Supporting Cognitive Processing in Multimedia Learning: The Use of Implementation Intentions

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. Kim Rouven Colin Stalbovs

aus Böblingen

Tübingen

2015

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

Tag der mündlichen Qualifikation:	21.01.2016
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Katharina Scheiter
2. Berichterstatter:	Prof. Dr. Caterina Gawrilow

Acknowledgements

My deepest gratitude goes to…

- Katharina Scheiter, who was a great mentor and supervisor all through my academic career and whose advice, support, and patience proved to be immeasurable and invaluable
- Caterina Gawrilow, who agreed to be the second reviewer of this dissertation.
- Thilo Poetzsch and my family for graciously supporting me at all times
- Halszka Jarodzka for sending me on paths that I would not have walked on my own
- Peter Gerjets for his support over the years
- All my friends and former colleagues at the *Leibniz-Institut für Wissensmedien* for their advice and feedback

Arahaṃ sammā-sambuddho bhagavā, buddhaṃ bhagavantaṃ abhivādemi.
Svākkhāto bhagavatā dhammo, dhammaṃ namassāmi.
Supaṭipanno bhagavato sāvaka-saṅgho, saṅghaṃ namāmi.

Table of Contents

1 Introduction

"Multimedia" has become a buzzword and catchall term for all kinds of colorful presentations of information, often suggesting an extensive use of state-of-the-art (digital) technology, adding pictures, sounds, animations, and/or film to regular text. The everyday use of the term often conflates the integration of different information channels, such as pictures, text, and video, with the use of digital devices to present such content, like computers, tablets, or smart phones. In turn, the term "multimedia" seems to have garnered an unwarranted specificity, becoming more associated with entertainment rather than (presumably) more serious practices like learning. Consequently, critics tend to question what added value multimedia could possibly have on the learning experience.

In the context of educational research, however, multimedia is more closely defined by the use of more than one code for the presentation of information (Mayer, 2009; Schnotz, 2002; Weidenmann, 2001). Codes are representational formats or sign systems for information and are processed in different but interacting channels of the cognitive system (Schnotz, 2002). In multimedia research, two types of codes are the most relevant: the verbal and pictorial codes. The verbal code represents language (words, sentences) and is a descriptive system of symbols, that is, a system of largely arbitrary signs that do not share similarity to the represented referents and whose meaning is established by convention (Clark & Paivio, 1991; Schnotz, 2002). The pictorial code is a depictive system of icons, that is, signs that somehow partake in the characteristics of the referent by means of visual characteristics (shape, color) or by means of spatial relations. Examples of the pictorial code can be found in pictures, animations, or films. The pictorial code can also refer to a system of symbols that conveys structural commonalities between reference and referent primarily by means of visual and spatial characteristics, such as found in graphs and diagrams (Schnotz, 2002). Thus, educational research usually defines multimedia as a combination of information in a verbal code and visual as well as spatial information in the pictorial code. Or to put it more simply, multimedia can be defined as a combination of written or narrated text and static or dynamic pictures. Thus, in the context of this dissertation, the term "multimedia" is used synonymously with the combination of (written) text and (static) pictures.

Conceptualized in this way, it can hardly be denied that multimedia finds ubiquitous use in textbooks, the Internet, as well as in other formal and informal learning resources. More importantly, it means that multimedia learning is far from being an innovation of recent modernity but instead has a long history, going back to when written texts started to use illustrations not just for the purpose of decoration but for conveying complementary information, for preventing misinterpretations by narrowing down possible interpretations of the text's meaning, or for helping to further elaborate the text's content (cf. Ainsworth, 1999). Such a use of illustrations can be found in works like illustrated versions of Dioscorides' pharmacopeia (e.g., the Vienna Dioscurides, ca. 515 CE), the agricultural treatises of Wang Zhen (1290-1333 CE), or the educational textbooks of John Amos Comenius (1592-1670 CE), to name a few. Thus, based on the fact that multimedia learning has been an everyday occurrence across the globe for a long time, the question of *whether* multimedia should be used at all becomes less interesting than the question of *how* multimedia should be used in the most effective way in order to support learning.

This being the case, research on the benefits of multimedia materials and the cognitive mechanisms underlying their processing has been highly prominent for the past decades. This research has primarily focused on ways of improving learning by optimizing the design of instructional materials (for an overview on multimedia research, see Anglin, Vaez, & Cunningham, 2004; Mayer, 2009). Such research focused on questions like the following exemplary list: Should text and pictures be presented consecutively or concurrently? Should they be presented in a separated or integrated fashion? Should text be narrated or written? Should pictures be static or dynamic? Are realistic or schematic illustrations more conductive to learning? All these questions aim at uncovering design criteria for multimedia materials that allow learners to cognitively process multimedia materials in the most efficient way, thereby optimizing learning.

The questions of how the design of the instructional material should be optimized to accommodate the processing capabilities of a learner's cognitive system, however, are only one part of the equation of multimedia learning. That is, how much learners benefit from multimedia may not only depend on the design of the learning materials but also on how skilled learners are in processing them (cf. Kombartzky, Plötzner, Schlag, & Metz, 2010). Therefore, the present dissertation aims at exploring the question how learning behavior can be modified to process multimedia materials in a more effective way.

As will be explained later (see Section 2 below), the quality of learning with text and pictures is dependent on the degree to which learners utilize effective cognitive processes, such as the selection, organization, and integration of text and picture information (Hegarty & Just, 1993; Mayer, 2009; Weinstein & Mayer, 1986). However, a number of studies have shown that learners often fail to make best use of the available learning materials on their own (e.g., Hannus & Hyönä, 1999; Kombartzky et al., 2010). Therefore, the series of four experiments presented in this dissertation investigated how learners can be supported in more effectively processing multimedia learning materials by relying on if-then plans (i.e., so called implementation intentions; Gollwitzer, 1999) to use effective cognitive processes. Even when the initiation of actions (e.g., the self-regulated application of effective cognitive processes during learning) presents itself as difficult, implementation intentions have been shown to facilitate an action's automatic initiation (Gollwitzer & Sheeran, 2006). This works even and especially under circumstances that make it difficult to initiate the action in the first place, such as high cognitive load or low motivation.

Experiment 1 investigated whether implementation intentions can foster the use of effective processing during multimedia learning, thereby improving learning. Furthermore, it tried to shed light on the question of how learners' task-specific motivation interacts with the use of implementation intentions. Experiment 2 focused again on the question whether implementation intentions can improve multimedia learning by supporting the underlying cognitive processes. Additionally, it addressed the question of what type of multimedia learning processes or combination thereof should best be supported by the use of implementation intentions. Experiment 3 aimed at replicating the main finding of Experiment 2 against a more conservative control condition. Finally, Experiment 4 tried to further delineate the differences between the use of implementation intentions and other effective ways to support the use of effective multimedia processes, more specifically the use of instructional prompts (e.g., Kombartzky et al., 2010; Thillmann, Künsting, Wirth & Leutner, 2009). In order to do so, the effect of both the use of implementation intentions and of prompts was studied under different conditions of cognitive load.

The dissertation is structured into three main parts: First, the theoretical background for the empirical part will be explicated in Section 2. Then, four experiments will be presented and discussed in Sections 3 to 6. Finally, the results of all four experiments will be discussed in Section 7.

2 Theory

2.1 Learning with Multimedia

One common and consistent finding in multimedia research is the so called ‘multimedia effect’ (Mayer, 2003). This effect describes that learning with illustrated text results in better recall and comprehension than learning with text alone (for overviews see Anglin et al., 2004; Fletcher & Tobias, 2005). A recent and fairly representative example of research concerning the multimedia effect is a study by Van Genuchten, Scheiter, and Schüler (2012): In order to investigate whether the multimedia effect varies in strength with different learning tasks, Van Genuchten and colleagues gave learners three types of learning tasks: conceptual tasks pertained to learning conceptual structures (e.g., relationships between people), causal tasks pertained to learning discrete cause-and-effect chains (e.g., the workings of a machine), and procedural tasks pertained to learning the temporal order and spatial relationships of actions (e.g., harvesting techniques). One group of participants learned with text alone, the other with a combination of text and pictures. The authors found a multimedia effect, that is, improved learning for the groups that learned with text and pictures, for all task types on a variety of learning measures. Yet, there were differences in the strength of the multimedia effect; it turned out to be stronger for procedural tasks than for conceptual or causal tasks. Although this study found the multimedia effect for all task types, the interaction between task types and the presentation type highlights that pictures can serve several functions when accompanying a text and thus may vary in their usefulness for learning.

In fact, the multimedia effect implicitly presupposes that pictures actually add to the text in some fashion instead of serving a purely decorative function. Based on a functional analysis by Ainsworth (1999), pictures can support text by either giving different information than the text or information that are very difficult to convey via text (e.g., spatial information), or by constraining a text’s interpretation by means of more specific information (e.g., displaying a specific color or shape that is not explicated in the text or narrowing down the meaning of an ambiguous word), or by stimulating the construction of a deeper understanding (e.g., by demonstrating how a mathematical equation translates into a specific diagram). When pictures are used in such a fashion to support a text, they can help learners to improve the recall and comprehension of the multimedia contents.

Ainsworth (1999) explains the advantage of multimedia presentations over monomedia presentations from a functional perspective, but what is the cognitive basis for the multimedia effect from an information processing perspective? Why can we learn a text better when it is illustrated? According to multimedia theories, such as the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2009) or the Integrated Model of Text and Picture Comprehension (ITPC; Schnotz, 2002; Schnotz, 2005), text-picture combinations offer an advantage over plain text because information can be extracted from both representational formats and can then be connected and integrated into a more comprehensive mental model. Learning with text and pictures first involves a number of processing steps, such as the extraction, processing, and organization of information, depending on the code-specific demands of each representational format (i.e., text processing, picture processing). Learners direct their attention to and select relevant information by identifying central words and sentences in the text or important components in the picture according to task-specific criteria. Based on grammatical features of the information extracted from the text, learners construct a surface structure of the text, which is then semantically processed and encoded in a propositional format. Equally, learners perceptually process the information in the picture in order to construct a visual image that serves as a basis for a pictorially encoded model of the picture's content. After having processed both representational formats individually, both theories of multimedia learning then assume a number of higher-order processes of building connections between the encoded information, effectively integrating the information into one coherent and comprehensive mental model of the information that are described and displayed in text and picture respectively (i.e., integration, Mayer, 2009; coherence formation, Seufert & Brünken, 2006). The pictorially encoded contents of the picture are cross-referenced and connected with prior knowledge and the propositional representation of the text's contents. Finally, all the connected information is integrated into a flexible mental model of the text's and picture's structure and content. Thus, the content of the text-picture combination is not only perceived and processed, but actually understood.

Yet, this depth of processing does not occur automatically. In fact, one of the core assumptions of the theories of multimedia learning (e.g., Mayer, 2009) is that meaningful learning with multimedia requires effort, that is, an active processing of information from both representational formats. Unfortunately, learners seem to have difficulties with fully engaging in this active information processing, as can be concluded from two lines of research.

First, eye-tracking studies have shown that if learners' processing of multimedia materials is unguided, they tend to focus on the text while neglecting the pictures (e.g., Hannus & Hyönä, 1999; Scheiter & Eitel, 2010; Schmidt-Weigand, Kohnert, & Glowalla, 2010a; Schmidt-Weigand, Kohnert, & Glowalla, 2010b). For instance, Schmidt-Weigand and colleagues (2010b) studied the effectiveness of an animated multimedia instruction about the formation of lightning, either with written or spoken text. At the same time, they recorded participants' eye movements during learning. The speed at which participants learned the 16-step instruction was self-paced. Schmidt-Weigand and colleagues found that participants in the spoken text group spent significantly more time looking at the visualization than in the written text group, a result that is expected when participants have only the visualization to look at while listening to the spoken text. However, participants in the written text group, who both had to read the text and look at the visualization, spent thrice as much time on reading the text as looking at the visualization. In this group, self-paced learning time highly correlated with time spent on reading the text, that is, additional learning time was used for reading, whereas the ratio between time spent on text and the visualization did not change with increased learning time. In this study, there were no differences in learning outcomes between both experimental groups; this result is not necessarily surprising, given that learners focused primarily on the text in the written text condition, whereas learners in the spoken text condition probably could not thoroughly and strategically process the text. Similarly, Hannus and Hyönä (1999) let school-aged children learn a biology lesson with an illustrated textbook while recording their eye movements. The children spent very little time on inspecting the illustrations. Only a mere 6% of the total learning time was spent on looking at the pictures. Yet, how creating strong connections between text and pictures can be helpful for learning demonstrates a second line of research.

This line of research investigates whether prompting or cueing learners to connect text and picture information results in better learning. A consistent finding in this field of research is that learners do benefit from such an instructional support (e.g., De Koning, Tabbers, Rikers, & Paas, 2009; Kombartzky et al., 2010; Scheiter & Eitel, 2010; Seufert, Brünken, & Zander, 2005). Seufert and colleagues (2005) investigated whether learners can be supported in building referential connections with small instructional design choices. They compared three groups that learned with a multimedia learning environment about the human circulatory system: One group received an illustrated written text; another group also learned with an illustrated written text but some words were hyperlinked so that, when learners moved their mouse cursor over the hyperlink, an arrow appeared and connected the word with the corresponding picture element; finally, one group learned with pictures and narrated text. While the group with narrated text had the best learning outcomes, amongst the two groups with written text, the hyperlinked text led to an improvement in learning outcomes. That is, the hyperlinks in the text increased learners' coherence formation when processing both representational formats, thereby increasing their comprehension. In a similar vein, Scheiter and Eitel (2010) let learners study an illustrated text about the human circulatory system.

Learners either received no further support (control group) or the instructional materials included signals (e.g., color-coding a word and its corresponding element in the picture, deictic references such as “see the arrows in the picture”, etc.) that marked corresponding elements in the text and pictures. During learning, learners’ eye movements were recorded. They found that the signals had an impact on learners’ gaze behavior and that, in turn, the gaze behavior acted as a mediator for learning outcomes. Thus, the signals helped learners to make better use of the picture, thereby fostering their learning. This finding therefore strongly suggests that learners seem to not make optimal use of the learning materials on their own.

Combining these findings indicates that learners often fail to take both representational formats, that is, text and pictures, fully into account. By failing to do so, learners do not utilize the additional and helpful information presented by the pictures (or the text, as the case may be), which in turn results in suboptimal learning. At the same time, it has been shown that the enrichment of learning materials with additional support for the integration of information from both representational formats can ameliorate learners’ deficient approaches to learning.

On the surface, these findings seem to suggest that we can solve the problem of suboptimal use of multimedia materials by just giving learners additional instructional support in the learning materials all the time. However, it should be remembered that the design of learning materials is only one part of the equation of learning; what remains are the capabilities of the individual learners to effectively use the learning materials that are provided to them. In fact, an overreliance on instructional guidance might actually prove suboptimal in the long run, as there is no guarantee that there will always be support for learners. Moreover, by giving learners maximal external support, they are unable to develop their own self-regulatory learning skills beyond the minimal level (Boekaerts, 1999). Thus, instead of only improving and enriching learning materials, it might also prove helpful to teach learners ways to process multimedia materials more effectively (Kombartzky et al., 2010; Pressley, Borkowski, & Schneider, 1989).

2.2 Effective processing of text and pictures

As explained above, theories of multimedia learning, like the CTML or the ITPC, propose a number of processes, such as the individual processing of the text up to the propositional level, the individual processing of the picture up to the pictorial level, as well as the higher-order coherence formation across both representational formats in order to create a rich and comprehensive mental model (or situation model; cf. Van Dijk & Kintsch, 1983). Thus, effective learning with multimedia relies on a balanced and active processing of text and pictures as well as the integration of information from both sources. This raises the question of which cognitive processes are important when learning with text-picture combinations and should thus be supported? While there are numerous helpful cognitive processes in multimedia learning, in this dissertation, I considered nine representative ones: three processes regarding the extraction and encoding of information from text (*text processes*), three processes regarding the extraction and encoding of information from pictures (*picture processes*), and three processes that are involved in connecting the text information with the picture information (*coherence formation* or *integration*). These processes were chosen based on both theoretical considerations like the theories of multimedia learning as well as based on empirical findings about what kinds of cognitive processes good learners tend to use in multimedia learning. Although these nine processes are hardly the only ones beneficial to multimedia learning, they constitute a good, representative sample of them.

2.2.1 Text processes

The three text processes described in this Section pertain to the extraction and processing of information from the text. Based on the extracted information learners create a surface structure of the text. This surface structure is then used to generate a propositionally encoded representation of the semantic content, a so called text base (Van Dijk & Kintsch, 1983). In the framework of both theories of multimedia learning (i.e., CTML and ITPC), these processes concern the selection and organization of information from the text.

One effective text process is the careful inspection of the text's headings on each page (*text overview / global text processing*). By first processing the overarching structure of topics and subtopics in a text, learners can construct a more hierarchical model of the text which, in turn, allows an easier access to memorized information in a top-down manner (Sanchez, Lorch, and Lorch, 2001). Research by Sanchez and colleagues (2001) indicated that the inclusion of headings in a text leads to an improved memory of the text's contents. Similarly, Hyönä, Lorch, and Kaakinen (2002) found in an eye-tracking study that readers who devoted more time to the topic structure of a text (i.e., headings) produced better summaries of the text in question.

Another effective text process is the careful rereading of all paragraphs after the first read-through of each page (*text rehearsal process*). In this context, rereading acts as a simple rehearsal and thus memorization strategy (O'Shea, Sindelar, & O'Shea, 1985; Weinstein & Mayer, 1986) but it also doubles as a monitoring help to make learners more aware of potential gaps in their understanding (Butler & Winne, 1995; Thiede, Anderson, & Therriault, 2003). When learners realize that some relationships described or implied in the text are not fully understood, they can return to the unclear text passages, reread it and possibly correct their lacking understanding. Moreover, it could be shown that rereading can act as a compensation for learners with inefficient verbal working memory capacity (Walczyk, Marsiglia, Johns, & Bryan, 2004).

Within the framework of the theories of multimedia learning, the global text processing as well as the text rehearsal process can be understood as generally supporting the selection and organization of information. One straight-forward and basal process pertaining to the organization of information from the text and thus of creating the text base is the connecting of information from one paragraph with information in previous paragraphs (*text* *organization process*; cf. Mayer, 2009). Text comprehension research has shown that this type of "filling in the blanks" (also called "bridging inferences") helps learners to form a more cohesive mental model of the text, thereby improving comprehension (e.g., McNamara, Levinstein, & Boonthum, 2004; Zwaan & Singer, 2003).

2.2.2 Picture processes

Analogously to the text processes, the three picture processes described in this Section broadly pertain to the selection and organization of information from the picture. They help learners to pay attention to, perceptually process, and encode information from the picture in order to create a visual model of the visual and spatial information in the picture.

Since learners tend to proceed in a rather text-driven way when processing multimedia materials (e.g., Schmidt-Weigand et al., 2010b), one effective picture process is looking at the pictures thoroughly before reading the accompanying text (*picture overview*). Eye-tracking studies have shown that initially glancing at a picture can help learners construct a coarse, holistic representation of the picture's visuospatial features before reading the text (Eitel, 2013; Eitel, Scheiter, & Schüler, 2012; Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013). Such a first impression can then act as a pictorial "scaffold" that guides the subsequent reading process and facilitates mental model construction (cf. Gyselinck & Tardieu, 1999). Within the framework of the theories of multimedia learning, this process can be viewed as supporting the selection and organization of picture information by allowing learners to get an early impression of relevant picture elements and their visuospatial relations, thereby forming a holistic representation of the picture information.

For more directly improving the selection of information in the picture (cf. Mayer, 2009), the search for crucial elements or components in a given picture constitutes an elemental cognitive process (*picture selection process*). The distinction and organization of different graphic elements in a picture is mostly directed by very basal routines (e.g., Gestalt laws; Wertheimer, 1938), but learners furthermore need to decompose the picture into smaller semantically meaningful components in order to infer each component's function in the picture's larger context (Hegarty & Sims, 1994).

These components can then be connected with each other. By linking the different picture elements and inferring how these elements interact with each other, learners can then construct a mental model of the content in a piecemeal fashion (Hegarty & Just, 1993; Hegarty & Sims, 1994). Hence, the creation of meaningful connections between all the relevant elements in a given picture, that is, organizing the information gained from the picture, represents another effective picture process (*picture organization process*; cf. Mayer, 2009).

2.2.3 Text-picture integration

Whereas the above mentioned six processes broadly concerned the selection and organization of information from their respective representational format according to the format's specific demands (i.e., text or picture processing), text-picture integration describes the higher-order process of generating referential links between all multimedia information, thereby constructing a comprehensive mental model or coherent mental representation of a material's intended meaning (cf. Kintsch, 1988; Mayer, 2009; Schnotz & Bannert, 2003; Zwaan & Radvansky, 1998). This mental model is not solely based on propositions gleaned from the learning materials; that is, it does not only include the basic semantic units of the text and the meaning of the picture, which learners constructed by mapping the picture's visuospatial relations onto semantic relations. It also interconnects the information in the text and picture with the learner's prior knowledge and experiences (Rouet, 2006; Schnotz, 2002). The resultant mental model is then incrementally updated by processing, interpreting, and integrating new information (Zwaan & Radvansky, 1998). In essence, by activating the individual elements in the text or picture in working memory and meaningfully relating them with each other and with prior knowledge, learners construct a mental model that organizes its information according to the described situation rather than the text or picture structure itself (Mayer, 2009; cf. Rouet, 2006; Van den Broek, 2012). The process of making sense of new information by drawing on prior knowledge in long-term memory is called inference generation and pertains to the “filling in the blanks” when relations between information are only implied. How well learners can generate inferences and thus process the learning materials more deeply depends both on individual factors, such as working memory capacity, but also on characteristics of the learning materials, such as the proximity of related information or the organization of information (cf. Lorch, Lemarié, & Grant, 2011; Van den Broek, 2012; Walczyk, Marsiglia, Bryan, & Naquin, 2001).

Two experiments by Hegarty and Just (1993) have shed some light on how learners integrate specifically multimedia information when learning about pulley systems. In the first experiment, they compared the learning outcomes of three groups that had learned either with text alone, pictures alone, or a combination of both. In this first study, they found the multimedia effect with regard to kinematic information (i.e., information relating to the motions of the pulley systems), so that the group with the text-picture combination learned better than the other two groups, whereas the groups learning with text or pictures alone did not differ in their learning outcomes amongst each other. In the second experiment, the authors used eye-tracking technology to have a closer look at how people learned from the text and pictures. They were interested in the gaze pattern of learners, for instance, whether learners would interrupt their reading regularly to look at the pictures or whether they would first read the text and then look at the pictures, or vice versa. In this particular experiment, learners interrupted the reading of the text several times; they inspected the picture on average six times and generally more often for complex pulley systems. Moreover, most inspections of the picture happened at the end of a sentence or idea unit (i.e., a unit of text that stated a singular configural or kinematic relation between pulley components). These results indicate not only a highly interleaved processing of the multimedia materials and construction of the mental model, it also shows that learners tend to first process and encode a meaningful unit of text before they inspect the picture. The latter result, in turn, suggests that the picture helps learners to check their understanding of and elaborate the relations between components instead of just memorizing the components themselves. With regard to the gaze data on the pictures, they identified two different types of gaze behavior: Local inspections were short and restricted to only one or few adjoined picture components, and were interpreted as the establishing of connections between text elements and picture elements. In fact, learners usually showed local inspections on picture elements that represented the content of the more recently read text unit. Global inspections, on the other hand, were generally longer and involved more than one or two adjoined components; they were interpreted to represents the construction of the mental model by combining the detailed mental representations created by local inspections. Learners showed more global inspections at the end of a text, implying that these kinds of inspections help learners to verify their comprehension of the picture and to integrate the information from the picture into their mental model of the multimedia materials. Overall, the results of Hegarty and Just (1993) suggest that leaners construct the mental model of the multimedia learning materials incrementally; they first read some sentences or a paragraph, integrate the information of the text at the level of text base, and then use the diagram to construct the mental model.

In a more recent eye-tracking study, Mason, Tornatora, and Pluchino (2013) could corroborate the importance of gaze transitions for integrative processes; they found that looking back and forth between a text and the accompanying picture was associated with better learning outcomes.

The first two processes investigated in this dissertation roughly correspond with local inspections (cf. Hegarty & Just, 1993) and concern the establishing of connections between semantically meaningful units of the text and the corresponding elements in the pictures. That is, the first process is the looking for the referred picture element after having read a text paragraph, whereas the second process is to look for the corresponding text paragraph after having inspected a picture element (*text-picture* and *picture-text integration processes*).

The third integration process roughly corresponds with global inspections (cf. Hegarty, 1992; Hegarty & Just, 1993) and is based on the finding that learners tend to study the picture after having finished the text, possibly to check their mental model of the depicted content and match it with their understanding of the text. Therefore, it is assumed that looking at a picture again before opening a new page, thereby verifying whether the picture matches one's own understanding of the text, constitutes an effective multimedia process (*matching of mental models*).

2.2.4 The relative importance of effective cognitive processes during multimedia learning

Supporting the nine, above-mentioned cognitive processes should result in better learning. Since the integration of information from text and pictures (i.e., the propositional representation of the text and the visual and partly propositional representation of the picture) plays an important role in incrementally constructing meaning from the multimedia content, it is expected that gains in learning are largest when there is support for integration processes; after all, without these integrative processes, the verbal and pictorial models are not connected in a meaningful way, so that there is hardly any added value from having available two different representational formats (Mayer, 2009). Nevertheless, learners need to construct a text base of the text *and* a perceptual image the picture as a prerequisite for a successful integration of text-picture information. As a consequence, when thinking about supporting effective cognitive processes during multimedia learning, the most promising approach seems to be a widespread support of different cognitive processes, that is, the concurrent support of text processing, picture processing, as well as text-picture integration (cf. Kombartzky et al., 2010).

2.2.5 Gaze data as measures of cognitive processes

As can be seen in the above-cited studies, one way to determine how effectively learners process multimedia materials is the assessment of eye movements by means of eye-tracking. By recording and registering the position and movement of a person’s gaze on a visual stimulus (e.g., pictures and/or text), eye-tracking gives insight into a person’s allocation of attention, thus revealing what parts of a stimulus the person paid attention to, in what order they attended different parts of the stimulus, and how long their gaze lingered on specific parts (Scheiter & Van Gog, 2009). Eye movements are generally divided into two types: saccades and fixations (Rayner, 1995). During saccades, the gaze moves quickly across the stimulus. This quick movement makes the extraction of information from the stimulus impossible (saccadic suppression). Instead information is extracted when a person’s gaze remains relatively still on one place during so called fixations.

According to Just and Carpenter (1980), the interpretation of information happens immediately at all levels of information processing (immediacy assumption), while there is no delay between fixating something with the gaze and processing it in one’s mind (eye-mind assumption). On basis of these two assumptions, gaze data can be used as valid indicators of underlying cognitive processes. Consequently, the assessment of gaze data represents an unintrusive way to obtain indicators of cognitive processes during learning (Johnson & Mayer, 2012; Mayer, 2010; Scheiter & van Gog, 2009; van Gog & Scheiter, 2010). Learners’ total fixation time on the text or the pictures indicates the degree to which they engage with the two representational formats; that is, the longer learners look at one representational format, the longer they can be assumed to process the information in it (cf. Johnson & Mayer, 2012). At the same time, looking back and forth between text and pictures has been shown to serve as a good indicator for learners' integration processes (e.g., Mason et al., 2013). In two of the four experiments that are presented in the present dissertation, learners' eye movements were used as a measure to investigate the questions how learners' gaze data change depending on which multimedia learning processes are supported and whether eye movements would predict learning outcomes in turn.

2.3 Causes of insufficient cognitive processing of text and pictures

Why do learners fail to sufficiently use effective cognitive processes when learning with multimedia materials? If learners are to be supported in a skillful and effective manner, it is important to accurately analyze at which point their learning process fails or gets derailed. According to Flavell, Beach, and Chinsky (1966), there are basically two reasons why learners fail to use effective learning processes: The first reason is the so called mediational deficiency, that is, that learners are actually incapable of utilizing these effective processes for some reason, for instance, because they are not cognitively mature enough for the cognitive process to result in the desired outcome or because they have never correctly learned the cognitive process in the first place. An appropriate instructional response for this type of deficiency is to teach learners the effective cognitive processes by means of specific trainings (cf. Kombartzky et al., 2010). Yet, even teaching these processes might not be enough to elicit them at the right time. The second reason for learners' insufficient cognitive processing suggested by Flavell and colleagues (1966) is the so called production deficiency, that is, learners do know what to do but, for some reason, fail to use effective cognitive processes at the right time. One reason for such a failure can be traced back to learners' capabilities in self-regulated learning.

2.3.1 Self-regulated learning

Self-regulated learning is a broad concept in educational psychology that can be roughly defined as the process of systematically activating and sustaining thoughts, actions, and emotions for the purpose of attaining a learning goal (Schunk & Zimmerman, 1994). More specifically, Boekaerts (1999) conceptualized self-regulated learning as a nested, three-layered construct encompassing interconnected aspects of how learners steer their learning processes, namely cognition, metacognition, and motivation (s. Figure 1). Thus, the construct's layers not only cover three elementary psychological constructs involved in learning, they also delineate different vectors of self-regulated learning (i.e., cognitive strategies, metacognitive knowledge and skills, or goals and resources) and the objects of the self-regulated processes (i.e., the learning content, cognition during learning, or the goal-directed learning action, respectively).

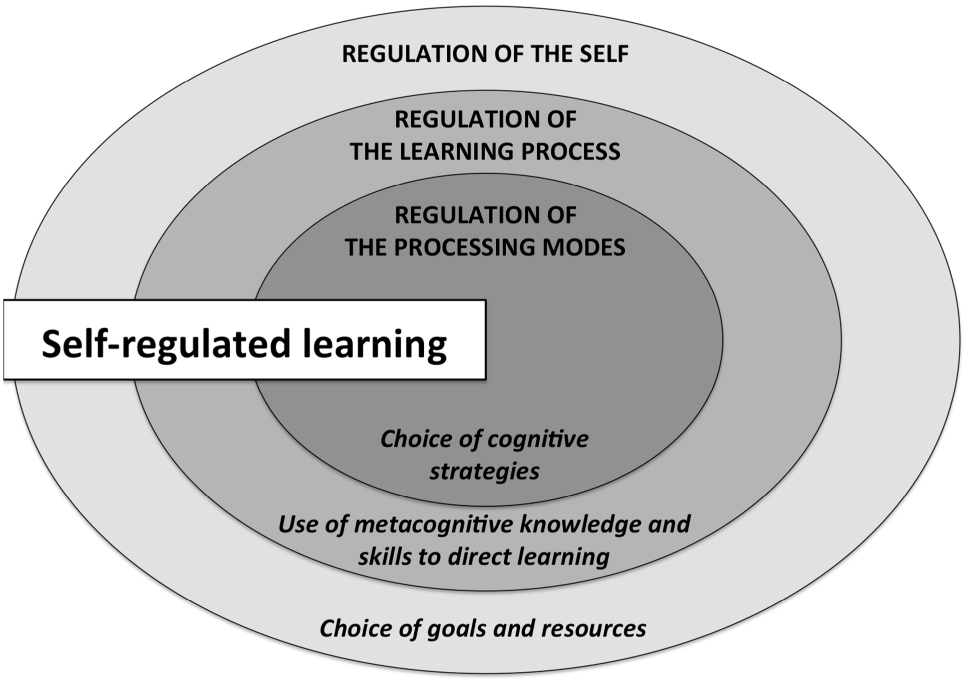


Figure 1. Boekaerts' (1999, p. 449) three-layered model of self-regulated learning.

At the most basic level, self-regulated learning concerns the regulation of processing modes, that is, learners' abilities in selecting, combining, and coordinating their cognitive strategies or processes in an effective way for learning (Boekaerts, 1999). This level encompasses basic strategies of learning, for instance rehearsal or elaboration processes (Weinstein & Mayer, 1989) or the multimedia processes described in Section 2.2. Furthermore, it encompasses emergant structural patterns of cognitive strategy use, such as surface-level or deep-level learning styles (Cassidy, 2004; Entwistle, Hanley, & Hounsell, 1979). At this level of resolution, the object of the learning processes is the learning content itself and thus self-regulation concerns *what* learners do with the learning materials, that is, what cognitive processes learners use in order to successfully extract and encode information from the learning materials.

The second layer of Boekaerts' model pertains to the metacognitive aspect of self-regulation, that is, the self-regulation of learning processes and learners' ability to direct their learning. In accordance with common definitions of metacognition (e.g., Veenman, Van Hout-Wolters, & Afflerbach, 2006), this level of self-regulation encompasses learners' metacognitive knowledge as well as their metacognitive skills. Metacognitive knowledge describes declarative knowledge about what types of cognitive and metacognitive strategies are appropriate in interaction with the task and learners themselves. Metacognitive skills describe learners' procedural knowledge about how to regulate their learning and problem solving activities, such as the planning, execution, monitoring, evaluation, and correction of cognition and behavior during learning (Boekaerts, 1999). Consequently, in this layer, the primary object of self-regulation is the innermost layer, that is, learners' cognitive processes during learning instead of the learning content itself.

Finally, the third and outermost layer involves learners' regulation of the self or their motivation and volition. On the one hand, this layer pertains to motivational factors, such as the choice of learning goals, the coordination between competing learning goals, or the allocation of resources and effort. On the other hand, it also involves *volitional* factors. Volition is a concept in motivational psychology that describes how goals or intentions are planned and executed and thus translated into action (Achtziger & Gollwitzer, 2008), for instance, by initiating actions, persevering even under difficult circumstances in the course of actions, or disengaging from actions (Boekaerts, 1999). Research in volition is especially concerned with the question of what problems can arise during the course of an action (e.g., failing to get started or getting distracted) and how such problems can be overcome. Essentially, this aspect of self-regulation concerns the question of *why* learners learn in certain ways, *why* they invest effort into learning under some conditions but not others, and *why* they sometimes fail to do what is expected of them. Moreover, it describes the motivational and volitional impetus that is intricately interconnected with the two inner layers of the model, which fuels and sustains the cognitive and metacognitive aspects of self-regulatory activities (Boekaerts, 1997), or that regulates non-cognitive internal processes like emotions (Corno & Kanfer, 1993). Thus, this layer has as a primary object the goal-directed learning action itself including its cognitive and metacognitive self-regulatory aspects.

Returning to the above mentioned production deficiency (Flavell et al., 1966), a failure to initiate and sustain effective cognitive processes during learning can then be traced back to any of the three layers of Boekaerts' (1999) self-regulation model. Therefore, Veenman and colleagues (2006) suggest that instructional support for improving self-regulated learning should follow the so called "*WWW&H* rule" (what to do, when, why, and how). That is, in order to offer a broad support for self-regulated learning, an instruction should address all three layers of self-regulation. Thus, supporting the *what, when,* and *how* covers both the cognitive and metacognitive aspects of self-regulated learning by teaching learners useful processes as well as how and when to use them in practice (e.g., by embedding strategy information in a learning environment in order to create a strong temporal contingency between processes and opportune moments for their use). If the instruction additionally explains *why* leaners should use certain processes, thereby increasing learners' effort and perseverance in learning, they also address the motivational part of self-regulated learning.

2.3.2 The role of motivation and volition in self-regulated learning

How exactly do the different layers of self-regulated learning interact? More specifically, how does motivation impact the cognitive and metacognitive processes involved in learning and thereby the quality of learning? Since motivation determines the setting of learning goals and, together with volition, represent the impetus behind learning activities (Boekaerts, 1997, 1999; Corno & Kanfer, 1993), such as the use of effective cognitive or metacognitive processes, it is also important to understand the mechanism behind this influence.

The level of engagement and effort required by the use of effective cognitive processing is more demanding than what students are generally used to (Pintrich, 1999). Consequently, a number of studies have found a positive relationship between motivational factors and the use of (meta-)cognitive processes (e.g., Pintrich, 1999; Pintrich & De Groot, 1990; Wolters, Yu & Pintrich, 1996). For instance, when learners find a task interesting, they are more likely to use deep processing strategies (Schiefele, 1991). Pokay and Blumenfeld (1990) analyzed to what degree different motivational factors predicted the use of effective learning processes over the course of a semester and found that, at the beginning of the semester, learners' perceived value of the subject matter and expectancies for success significantly predicted learners' use of cognitive strategies, whereas value also predicted the use of cognitive strategies later in the semester. Berger and Karabenick (2011) found that a higher self-efficacy predicted a more frequent use of deep processing strategies, while a higher value predicted a more frequent use of rehearsal strategies. In effect, learners who are less motivated with regard to the learning task seem to encounter a lot more difficulties when using effective cognitive processes than learners who are more motivated concerning the task at hand.

Theory

The Cognitive-Motivational Process Model by Vollmeyer and Rheinberg (2006) is one model in motivational research that attempts to explain the continuous influence of motivation on self-regulated learning. According to the model, the use of effective learning processes is one of several mediating factors between a learner's initial motivation and their learning performance. Learner's initial motivation is regarded as task-specific and is defined in terms of four factors: (1) probability of success, (2) interest, (3) challenge, and (4) anxiety. The first of these factors, probability of success, is based on the assumption that learners implicitly calculate the likelihood of their success concerning the task at hand, taking into account their own perceived abilities and the perceived difficulty of the task. Anxiety, on the other hand, is understood as learners' fear of failure in a specific situation; fear of failure, in turn, has been shown to have a negative impact on learners' metacognitive self-regulation (Bartels & Magun-Jackson, 2009). Interest represents the subjective value that learners attribute to the task. Finally, challenge represents how much learners accept the task as an achievement situation that they intend to succeed in; as such, it also represents the importance that learners assign to the task. Moreover, learners need personal achievement standards with which they can compare their performance in order to interpret a task as challenging (Vollmeyer & Rheinberg, 2000). In one of their studies, Vollmeyer and Rheinberg (2006) aggregated probability of success, interest, and challenge as a variable for initial motivation and found that learners with high initial motivation used more systematic strategies in a self-regulated learning task and reported a higher motivation during learning.

While the Cognitive-Motivational Process Model makes assumptions about how the use of effective cognitive processing act as mediators between motivation and learning outcomes, it still leaves the question unanswered of how motivation influences the use of effective cognitive processes exactly.

Corno (1986; see also Corno & Kanfer, 1993) conceptualizes learners' failure to use effective learning processes as a problem of volition. One prominent volitional theory is the

Rubicon model of action phases (Heckhausen & Gollwitzer, 1987). The Rubicon model defines actions as “all activities directed toward an ‘intended goal’” (Achtziger & Gollwitzer, 2008, p. 272), and divides the natural course of an action into four functionally distinct phases: (1) the predecisional phase, in which goals are set based on wishes and considerations of an action’s probability of success, (2) the preactional phase, in which concrete strategies are planned for realizing the goal, (3) the actional phase, in which the planned strategies are implemented, and (4) the postactional phase, in which the outcome of the goal striving process is evaluated. According to the model, the first and the last phase of an action are considered to be motivational in nature in that they concern the setting of goals and evaluation of behavior, whereas the two phases in the middle are considered to be volitional in nature. Based on this framework, the general choice of processing strategies, whether automated or deliberate, is assumed to happen in the preactional stage, while the actual and effortful use of effective cognitive processes is assumed to happen in the actional phase. Since all of the action phases are interdependent, the degree to which effective cognitive processes are initiated in the actional phase depends on the volitional drive that has been built up in the phases beforehand. The drive underlying an action is called the *volitional strength* and determines how much effort is invested in seeing an action to its end. The volitional strength is largely determined by early motivational deliberations prior to goal setting, such as the probability of success and the personal value attached to a goal, as well as the commitment to the goal. In this way, the Rubicon model explains how a high initial motivation will lead to a more consistent and persistent use of deep level processing during learning, while low initial motivation can lead a learner to abandon or neglect the initiation of effective cognitive processes. In explaining the processes underlying goal-directed behavior in this way, the Rubicon model fulfills a similar function as other theories of self-regulation do in educational research (Corno & Kanfer, 1993; Wolters, 2003).

The Rubicon model can also explain the finding that the intention to achieve a goal does not necessarily lead to goal achievement (Sheeran, 2002). Several hindering problems can arise during the course of an action, such as failing to get started with goal striving, getting derailed during goal striving, failing to realize when to stop, or overextending oneself by pursuing several goals concurrently (Gollwitzer & Sheeran, 2006). In order to really translate an intention into action (e.g., integrating text and picture when learning with multimedia), the intention must either be undergirded by the necessary volitional strength, or the translation process must be supported by helpful self-regulation strategies.

2.4 Supporting learners' cognitive processing of multimedia materials

Would it be sufficient to simply inform learners about these effective cognitive processes so that they would make optimal use of multimedia learning materials? Based on prior research, this seems unlikely. In order to successfully process the learning materials effectively, learners do not only have to know what to do and how to do it, they also have to do it at the right moment (Pressley, Borkowski, & Schneider, 1987; Veenman et al., 2006). As mentioned before, despite knowing what cognitive processes would be beneficial for learning, learners often fail to effectively process learning materials on their own (cf. Flavell et al., 1966; Flavell, 1979; Winne, 1996). This failure can be caused by a number of factors, such as learners' unsuccessful monitoring of their own cognitive processes, the insufficiently automatized use of effective cognitive processes, or insufficient cognitive resources to rein in ineffective cognitive processes and initiate effective cognitive processes (Boekarts, 1997; Pressley et al., 1989; Winne, 1996; Wirth, 2009). Since using effective cognitive processes at the right time during learning has been proven to be difficult, educational research has tried to tackle the question of how learners can be supported in doing so. While trainings of effective cognitive and metacognitive processes have been shown to work, they are also very time-consuming (Hattie, Biggs, & Purdie, 1996). Moreover, they often primarily address what learners learning processes should use (i.e., the mediational deficiency) instead of when and how to use them (i.e., the production deficiency). One prominent, less time-consuming approach is the use of instructional prompts (Bannert, 2003, 2006).

2.4.1 Instructional prompts

Instructional prompts are generally defined as recall or performance aids for the induction and stimulation of activities relevant to the achievement of an educational objective such as cognitive, metacognitive, motivational, volitional, and/or cooperative learning activities (Bannert, 2009). That is, instructional prompts either help learners to recall concepts or procedures that support the learning process or they directly aim at activating effective learning processes. There is no established, standard presentation format of instructional prompts; for instance, they could consist of general questions that prompt learners to think about what study activities they ought to employ for a given task or about the rationale behind their problem solving steps. Alternatively, instructional prompts could consist of very specific step-by-step instructions on what learners should do in the course of learning.

Although instructional prompts usually do contain information about which learning processes are effective, their primary purpose is not to teach these processes to learners but rather to guide learners' attention to these helpful processes by supporting their recall and initiation. That is, they do not primarily address the mediational deficiency but the production deficiency.

A technologically simple example for the use of instructional prompts in multimedia learning can be found in Kombartzky and colleagues (2010). In their study, two groups of sixth grade students used an animation with narrated text in order to study the biology subject of "honey bee dances". One group of students watched the animation first and then had to write an essay on what they had learned (control group). Another group of students received a worksheet with written step-by-step instructions on what to do during watching the animation (prompts condition); they worked through the worksheet while watching the animation. The results of a post-test with three measures for learning outcome showed that, compared with the control group, the prompts group showed significantly better learning outcomes across the board.

A technologically more elaborate example for instructional prompts represents the study by Thillmann and colleagues (2009), in which ninth graders used a computer-based learning environment about the physics domain “buoyancy in fluids” that allowed them to plan and run simulated experiments. All students were prompted to generate data by running simulated experiments in which only one variable was manipulated and to document the relationship between variables. However, there were three groups that received the prompts at different times in the course of learning. One group received the prompts before learning with the computer-based learning environment, another group received the prompts during learning in a suboptimal order (they received the prompt to document their results first and the prompt to generate data second), and a third group received the prompts during learning in an optimal order (i.e., the prompt to generate data first, the prompt to document their results second). Thillmann and colleagues not only assessed learners’ conceptual knowledge before and after the learning phase, they also used logfiles to analyze whether learners actually showed the prompted learning behavior. The results showed that all learners showed an increase in conceptual knowledge after learning with the computer-based learning environment. Moreover, the results revealed that the two groups that were prompted during learning significantly outperformed the group that was prompted before learning. However, there was no difference between the two groups that received the instructional prompts in the course of learning. Finally, Thillmann and colleagues could show that the effect of prompts on learning outcomes was mediated by learners’ use of the prompted learning processes.

How and why do prompts work? On a surface level, instructional prompts follow the above mentioned "*WWW&H* rule" (Veenman et al., 2006); that is, they tell learners *what* to do (and optimally, *why*), *how* to do it, and in some cases (cf. Thillmann et al, 2009), *when* to do it. Hence, instructional prompts are able to offer a broad support for self-regulated learning, addressing cognitive and metacognitive (*what*, *when*, and *how*), as well as motivational (*why*) aspects of self-regulation. On a deeper level, however, they also address a problem of limited cognitive resources. The juggling of learning processes, like memorizing or organizing information, and motivational and metacognitive processes, such as setting goals, planning and monitoring the learning process, is cognitively demanding, especially for learners with little prior knowledge (Van Merriënboer & Sluijsmans, 2009; Wirth, 2009). Since learners' working memory capacity is limited (cf. Baddeley, 1992), these cognitive demands are accompanied by the constant risk of cognitive overload. Cognitive overload happens when the limited capacity of learners' working memory is exceeded by the requirements of the learning situation (Sweller, Van Merriënboer, & Paas, 1998). Learners' working memory capacity can be strained by the inherent difficulty of the learning task ("intrinsic cognitive load"), by impairing external and internal factors ("extraneous cognitive load"), such as badly designed learning materials, distractions, or ineffective processing strategies, or even by effective cognitive processes ("germane cognitive load"). When working memory capacity is too strained by the difficulty of the learning task or unproductive, extraneous factors, learners are not able to successfully encode the information in working memory and integrate it with prior knowledge anymore, thus resulting in decreased learning. Consequently, learners who are cognitively overwhelmed, will not be able to use deep-level learning processes (Van Merriënboer & Sluijsmans, 2009). In cases like these, instructional prompts are supposed to serve in the function of external compensators when learners fail to spontaneously initiate effective processes on their own, such as under conditions of high cognitive load. Accordingly, the findings of Thillmann and colleagues (2009) show that instructional prompts are most effective if they are given when the risk of cognitive overload is especially high, that is, during learning rather than before learning. Yet, even when instructional prompts are given in cognitive demanding situations and successfully elicit the initiation of effective cognitive processes, it is not guaranteed that these processes are fully executed and/or kept up; the conscious execution and upkeep of the effective learning processes still add to the cognitive load of the learning task, thereby possibly resulting in overload.

Still, instructional prompts have been shown to work in a wide variety of contexts (e.g., Davis, 2003; Lin & Lehman, 1999; Meijer & Riemersma, 2002; Schmidt-Weigand, Hänze, & Wodzinski, 2009). For instance, Chi, DeLeeuw, Chiu, and LaVancher (1994) prompted learners to elaborate an expository text by explaining the text contents to themselves and found that prompted learners gained a deeper level of understanding than learners who were not prompted. Such self-explanation prompts have also been shown to be effective for learning with worked examples, that is, examples that help learners with problem solving by guiding them step-by-step through the solution (e.g., Atkinson, Renkl, & Merrill, 2003; Reisslein, Atkinson, Seeling, & Reisslein, 2005). In a study by Azevedo, Moos, Greene, Winters, and Cromley (2008), a human tutor prompted some of the participants to activate their prior knowledge, to plan and monitor their learning activities, as well as to use effective cognitive processes during learning with a hypermedia environment. A control group had to self-regulate their learning without the help of prompts. Results of this study showed that prompted learners displayed more effective self-regulatory processes and consequently developed a better understanding of the learning content.

There have been a few studies that could show prompts to be effective specifically in the context of multimedia learning. For instance, Reinking, Hayes, and McEneaney (1988) investigated whether general and specific instructional prompts can guide learners' attention towards the graphics that accompanied the text. Their study revealed that instructional prompts helped to improve comprehension of the learning materials, especially for poor readers, whereas they had less of an impact on good readers. In a study by Weidenmann (1988), learners studied a text about psychology. In a control group, learners only received the text; in two other groups, they learned with either simple text or an illustrated text and received an instruction to relate the text's contents to their own experience; and in a prompt group, learners received the illustrated text with an instructional prompt to inspect the picture carefully and to relate the contents of the text to the picture. Weidenmann found that the group that received instructional prompts showed the best learning outcomes, whereas the other groups did not differ. Peeck (1994) compared three groups learning from a biology text: one group learned with the text alone (control condition), one with an illustrated text (multimedia condition), and one group learned with an illustrated text and the explicit instructional prompt to match the content of the text with the pictures (prompt condition). Although Peeck found no multimedia effect in this study, that is, the multimedia group did not outperform the control group, they found a significant increase of learning outcomes in the prompt condition. Finally, as already explicated above, Kombartzky and colleagues (2010) found that instructional prompts led to an improvement in three measures of learning outcomes.

However, not all studies gave evidence for the effectiveness of prompts in multimedia learning, or only under certain conditions. Hayes and Readance (1983) gave unspecific prompts to inspect the picture and found no positive effect on learning outcomes. Drewniak (1992) compared three different types of instructional prompt in a computer-based learning environment: Whenever learners looked at a picture in the learning environment, they were either prompted to monitor their understanding of the picture (metacognitive prompt), to memorize the picture information and to use the information as a structuring device when reading the text (integrative prompt), or direct their attention to specific and important parts of the picture (selective attention prompt). Additionally, there was a control group that received no prompts when looking at a picture. However, in her study, Drewniak found no difference in learning outcomes between the four groups. Brünken, Seufert, and Zander (2005) let learners study with a computer-based learning environment about the circulatory system. They had three groups in their experiment: a group that received pictures with written text, another group received pictures with narrated text, and a third group received pictures with written text and instructional prompts supporting coherence formation. The instructional prompt was a multiple-choice comprehension question that had to be answered whenever learners wanted to open a new learning page. They found that there was no difference between groups in learning outcomes relating to textual information. With regard to pictorial information, the results only revealed a superiority of the condition with narrated text, whereas there was no difference between the condition with written texts and the instructional prompts. Finally, Bartholomé and Bromme (2009) investigated how learners' coherence formation could be supported when learning with a computer-based multimedia learning environment about botany. They gave learners two types of cues for supporting the mapping of text and pictures (numerical labels and highlighting) and either gave no instructional prompts or presented a text box on each page of the learning environment that prompted learners to systematically inspect the picture and to relate the picture to the text. There was no main effect of prompts on learning; instead, prompts were effective only when combined with another instructional support measure (highlighting).

The reason for the mixed results of research on instructional prompts in multimedia learning might be twofold: First, prompts do not necessarily convey all the required information about how to initiate effective cognitive processes, such as information about when, why, and how to do it (e.g., Brünken et al., 2005; cf. *WWW&H* rule by Veenman et al., 2006). Second, the processing of prompts may put actually further cognitive demands on learners, thereby possibly causing cognitive overload (Bartholomé & Bromme, 2009), especially if learners receive the prompt only before learning and then have to keep the prompt in mind all the time during learning. As Thillmann and colleagues (2009) demonstrated, the effect of instructional prompts was stronger when the prompts were presented during learning rather than before it. Yet, even when instructional prompts reduce some amount of cognitive load due to self-regulation by being presented during learning, there is no guarantee that instructional prompts will work. For one, the conscious execution and upkeep of the prompted learning action also represents a form of cognitive load and can thus lead to cognitive overload when the learning task or learning environment are sufficiently difficult by themselves. For instance, Horz, Winter, and Fries (2009) gave undergraduate and graduate students a complex and authentic computer-based learning environment about cost and sales accounting. For some students, the learning environment contained situated instructional prompts, asking them to do optional tasks and look for additional information; for other students, it did not contain any instructional prompts. The authors also measured cognitive load during learning. Contrary to their expectations, the novice undergraduate students who were expected to benefit the most from the instructional prompts showed worse learning outcomes when they received the prompts, likely due to cognitive overload. Instead, the more experienced graduate students slightly benefited from the instructional prompts. Thus, learners with low prior knowledge who were in need of the instructional support were hampered by the additional cognitive load imposed by the situated instructional prompts. This means that, especially when learners are already at the limits of their cognitive capacities, the effectiveness of instructional prompts might actually disappear.

To summarize: since learners often fail to process multimedia materials effectively, due insufficient self-regulation, cognitive overload, or low motivation, they require additional support. One prominent way to provide such support is the use of instructional prompts which guide learners to use effective learning processes. Instructional prompts have been shown to work in many contexts, although the results in the context of multimedia learning have been mixed so far. One reason for the mixed results might be that instructional prompts do not contain all necessary information to use effective cognitive processes at the right time. Another reason might be that they do not reduce the cognitive load of difficult learning tasks enough (e.g., the cognitive load by the conscious execution and upkeep of the prompted behavior) or even put additional cognitive demands on learners (e.g., novice learners), thereby causing cognitive overload under cognitively straining circumstances.

Through the lens of self-regulation, learners' deficit to adequately use effective cognitive processes during learning with multimedia learning materials can be also conceptualized as a failure to successfully translate the goal to learn into specific behavioral responses that are supportive for attaining this goal (Corno, 2001). From this perspective, learners' difficulties in effectively processing multimedia materials mirrors the finding from self-regulation research that having an intention to do something (e.g., "*I want to successfully learn from this multimedia learning materials!*") does not automatically mean people will act upon these intentions (Sheeran, 2002). One self-regulation technique that has consistently shown to be effective in bridging this gap between intention and behavior is the use of implementation intentions (Gollwitzer, 1999).

2.4.2 Implementation intentions

Implementation intentions are specific "if-then" plans that aim at facilitating the translation of a goal-directed intention into actual action(s) (Gollwitzer & Sheeran, 2006). They achieve this by strongly linking a situational cue in the "if" part that indicates a good opportunity to act (i.e., a conditional trigger), with an action in the "then" part that is conductive for attaining the goal (i.e., a behavioral response). In contrast to simple goal intentions which only specify the goal to be achieved, that is, the *what* of an intended action, implementation intentions also specify the *when*, *where*, and *how* for achieving the goal

(Gollwitzer & Sheeran, 2006). Structurally, an example for an implementation intention in the context of multimedia learning could be: "***If*** *I have read a paragraph,* ***then*** *I will relate its information to the referred elements in the picture!*"

A typical example for a study investigating implementation intentions can be found in Achtziger, Gollwitzer and Sheeran (2008). They asked participants to halve their consumption of an unhealthy snack food of their choice in the week following the intervention. Half the participants answered a questionnaire measuring the strength of their intention to reach this goal, whereas the other half answered the same questionnaire and were then asked to internalize an implementation intention about their goal. Their implementation intention instruction read as follows (p. 384, Achtziger et al., 2008): "Please tell yourself: 'And if I think about my chosen food, then I will ignore that thought!'" One week later, the participants reported how many unhealthy snacks they had consumed in the past week. All participants had reduced their intake of unhealthy food but the reduction of the implementation intention group turned out to be significantly higher. Moreover, participants who had internalized an implementation intention had actually achieved their goal of halving their intake of unhealthy snacks, whereas the goal intention group had not.

The effectiveness of implementation intentions has been shown in meta-analyses and reviews incorporating studies with a wide range of samples, settings (laboratory, field), domains (e.g., problem solving, stereotyping, medication intake, cancer prevention screenings, dieting, exercising) and dependent variables (Gollwitzer & Sheeran, 2006; Koestner, Lekes, Powers, & Chicoine, 2002; Sheeran, 2002). A meta-analysis incorporating 63 studies indicated that implementation intentions have a medium to large effect (d = .65) on goal attainment (Gollwitzer & Sheeran, 2006). Beyond showing that implementation intentions are effective in supporting goal-oriented action, research has also provided evidence for the mechanism underlying this effect.

Theory

Two processes, which are linked to either the "if" or the "then" component, contribute to their effectiveness: First, when an implementation intention is formed, the situational cue in the "if" part becomes highly activated and thus accessible in memory. This accessibility then facilitates the recognition of suitable opportunities for initiating the intended action (Parks–Stamm, Gollwitzer, & Oettingen, 2007; Webb & Sheeran, 2008). For instance, Aarts, Dijksterhuis, and Midden (1999) found that, in a lexical decision task, participants who had internalized an implementation intention responded more quickly to words that were relevant to the anticipated situational cues. Even in more ecologically valid constellations, such as when goals were set by participants themselves, implementation intentions resulted in more attention to the situational cues and better recall of the situational cues (Achtziger, Bayer, & Gollwitzer, 2012). Thus, situational contexts that match the conditional trigger will be detected faster and more reliably as well as better discriminated from other stimuli. According to Gollwitzer and Sheeran (2006), studies investigating this part of the process underlying implementation intentions yielded evidence with an average effect size of d = .80 in favor of the process described above.

The second sub-process contributing to the effectiveness of implementation intentions concerns their "then"-part. Implementation intentions create a firm association between the situational cue and the intended action (Webb & Sheeran, 2008). Once the conditional trigger has been recognized, it will activate the implementation intention in memory, thereby initiating the intended action (e.g., Parks–Stamm et al., 2007; Webb & Sheeran, 2008). Due to this simple but strong cognitive mechanism actions that are evoked by implementation intentions share similarities with automatized behavior (Bargh, 1994; Gollwitzer & Sheeran, 2006). That is, these actions follow more swiftly on the situational cue (e.g., Gollwitzer & Brandstätter, 1997), they are more efficient with regard to cognitive resources (e.g., Brandstätter, Lengfelder, & Gollwitzer, 2001), and they are triggered even if the intention is not conscious at the time (e.g., Ajzen, Czasch, & Flood, 2009). Importantly, according to the review by Gollwitzer and Sheeran (2006), despite the simplicity by which implementation intentions can be induced in participants, they have long-lasting effects over weeks and months, thereby extending well beyond the situation in which the implementation intention has been initially internalized. Gollwitzer and Sheeran (2006) showed that all three features of automaticity evident in implementation intentions are associated with large effect sizes (immediacy: $d = .77$; efficiency: $d = .85$; lack of intent: $d = .72$). Therefore, implementation intentions allow the delegation of action control from the self to specific situational conditions, effectively creating instant habits (Gollwitzer, 1999).

Consequently, implementation intentions have been proven advantageous even under volitionally difficult conditions. For instance, Brandstätter and colleagues (2001) asked opiate addicts to write their curriculum vitae (CV) within a specified period of time. Some of these patients were under withdrawal at the time of the study, while some were already past their drug withdrawal. The authors chose patients with withdrawal symptoms because drug withdrawal is characterized by a strong and disruptive cognitive preoccupation that results in high cognitive load. After all patients had agreed to participate in the study and to write their CV, that is, after all of them had formed a goal intention to complete the task, they were instructed to internalize an implementation intention. One group was instructed to phrase, write down, and internalize an implementation intention about where and when they wanted to have lunch (irrelevant implementation intention), whereas the other group was instructed to phrase, write down, and internalize an implementation intention about where and when to write their CV (relevant implementation intention). The results showed that both withdrawal and post-withdrawal patients significantly benefited from having internalized a relevant implementation intention: They were more likely to write the CV and more likely to write it at the designated time. Moreover, patients under withdrawal benefited even more from the implementation intention than the post-withdrawal patients. In fact, they were the most likely to turn in the CV. Thus, the study demonstrated that the internalization of an implementation intention can significantly facilitate patients' translation of their goal intentions into action, even under circumstances characterized by high cognitive load. Gawrilow and Gollwitzer (2008) could show that children with attention deficit/hyperactivity disorder (ADHD), that is, a disorder that is associated with problems in action control, showed better inhibition in a task requiring a high level of executive control when they used implementation intentions than those who did not. Gawrilow, Morgenroth, Schultz, Oettingen, and Gollwitzer (2013) found that an intervention combining implementation intentions with mental contrasting (i.e., imagining the optimal outcome of a goal-directed action) significantly improved self-regulation in school children. Furthermore, this effect was even stronger for children at risk for ADHD.

These results suggest that implementation intentions could serve as a good support for the use of helpful cognitive processes, even under circumstances in which learners usually do not have a good control of action. Due to the compensatory effect of implementation intentions, they should not only work successfully under conditions of high cognitive load but also might compensate for other factors impairing self-regulated learning, such as a suboptimal motivational orientation. More specifically, implementation intentions should compensate for low task-specific motivation and a resulting lack of volitional strength to initiate and keep up effective learning processes. Based on these features and on the effectiveness of implementation intentions in a wide variety of settings, they seem like a promising support for the use of effective cognitive processing in multimedia learning.

Implementation intentions have already been shown to be effective in an educational context. An example can be found in Oettingen, Hönig, and Gollwitzer (2000): they compared two groups of students, one using an implementation intention and the other only forming a goal intention to do as many arithmetic tasks as possible, and found that the group using the implementation intentions showed far greater perseverance in doing arithmetic tasks than the group who had only the intention to do so. Bayer and Gollwitzer (2007) conducted two studies in which self-efficacy strengthening implementation intentions were investigated. In the first study, they let high school students perform a math task. One half of the students formed a goal intention to solve as many math problems as possible ("I will correctly solve as many problems as possible!"; p. 6) and the other half formed an implementation intention ("And if I start a new problem, then I will tell myself: I can solve it!"; p. 6) and memorized it for three minutes. Consecutively, the implementation intention group solved significant more math problems than the goal intention group. Duckworth, Grant, Loew, Oettingen, & Gollwitzer (2011) compared two groups of high school students who were preparing for a qualification exam during the summer break: One group of students received a short intervention with implementation intentions plus mental contrasting about completing as many practice questions as possible over the summer break, whereas the control group wrote a short essay about an influential person or event in their life instead. At the end of the summer break, the implementation intention group had completed over 60% more practice questions than the control group. Finally, in an example of implementation intentions in a quasi-educational context, Wieber, von Suchodoletz, Heikamp, Trommsdorff, and Gollwitzer (2011) investigated how well school-aged children could shield their efforts in a classification task from distractions of varying attractiveness. Implementation intentions helped children to shield their efforts from distractions of low, moderate, and high attractiveness, whereas children who had the intention to ignore distractions could shield themselves only from distractions of low attractiveness.

As can be seen from these examples, these studies focused on a diversity of factors during learning. To my knowledge, however, there has been no research so far regarding the question whether implementation intentions can support multimedia processing and thus improve learning. Hence, this dissertation attempted to address this research question by using implementation intentions as a means to foster the use of effective cognitive processes in multimedia learning.

A widespread support of cognitive processes is only possible if more than a singular implementation intention is used, however. Therefore, a second question strongly related to the first one is how many implementation intentions can and should be used concurrently. Generally, it has been shown that more than one implementation intention can be used successfully. For example, Achtziger and colleagues (2008) used implementation intentions as a means to control difficult inner states (e.g., anxiety) during physical exercise. They instructed participants to generate four personalized implementation intentions based on a list of negative states and coping responses. The findings indicated that participants who had used implementation intentions performed better during their physical exercise than a no-goal or goal intention group. De Vet, Oenema, and Brug (2011) investigated whether the specificity (concerning the "when, where, and how") and number of implementation intentions had an impact on participants' level of physical exercise. They found that the specificity of implementation intentions was a significant predictor for the effect of implementation intentions, while their number was not. However, they also found an interaction indicating that more implementation intentions are more effective only if they are highly specific at the same time. Since the cognitive processes underlying multimedia learning and, more specifically, those nine processes investigated in this dissertation are rather specific, it can be assumed that a larger number of implementation intentions should lead to better cognitive processing and better comprehension. Moreover, a larger number of implementation intentions should generally allow for a more flexible and widespread use of effective multimedia processes. On the other hand, several implementation intentions might interfere with each other, especially if the situational triggers in the "if" part are too similar or follow too closely on each other. Thus, learners might initiate an effective multimedia process just to abort it as soon as the situational trigger of another, similar implementation intention is encountered.

To summarize: learners often fail to effectively process multimedia materials. This deficit can be conceptualized as a failure to successfully translate the learning goal into specific behavioral responses for attaining this goal. One technique that has been proven to support this translation process is the use of implementation intentions. Implementation intentions represent specific if-then plans that strongly link cues for good opportunities to act with those behavioral responses that are effective for achieving a goal, thereby increasing the likelihood that the response is actually initiated. The specific features of implementation intentions, such as a high accessibility of the situational cue and the automaticity of the behavioral response, as well as their effectiveness in a wide variety of contexts suggest that they might be a good means to support the use of effective cognitive processing during multimedia learning. Three questions that warrant research are: Can implementation intentions support effective multimedia processing? What type of cognitive processes should be supported via implementation intentions? How many implementation intentions can and should be used concurrently?

2.4.3 Differentiating implementation intentions and instructional prompts

How do implementation intentions relate to instructional prompts? On the surface, they seem very similar: Both instructional prompts and implementation intentions contain information about “what” effective cognitive processes learners are supposed to initiate.

Regarding the “when”, implementation intentions have two advantages over instructional prompts. First, implementation intentions necessarily contain information about when to initiate effective processes in their “if” component; instructional prompts, on the other hand, can but do not have to contain this information. Second, in constrast to instructional prompts, implementation intentions make the trigger conditions in the “if” part cognitively very accessible (Parks–Stamm et al., 2007), so that these opportune moments to act are less likely to be overlooked. Thus, implementation intentions should facilitate effective cognitive processing at just the right time. With instructional prompts, it is much more difficult to precisely time their presentation.

Concerning the "how", the effectiveness of instructional prompts or implementation intentions depends on their specificity. Assuming that the cognitive processes supported by implementation intentions or instructional prompts are sufficiently specific, the question of "how" is less problematic than for vague, general processes that offer many ways for translating them into actual behavior. Nevertheless, it could be argued that, since there is a strong associative link between the "if" and the "then" parts in implementation intentions (Webb & Sheeran, 2008), this link should give implementation intentions an advantage over instructional prompts. After all, instructional prompts leave the initiation of effective processing more strongly to learners' self-regulatory capacity than to circumstance. That is, in this regard, instructional prompts share more similarities with simple goal intentions than with implementation intentions.

With regard to the "why", both instructional prompts and implementation intentions can but do not have to contain information about "why" effective cognitive processes should be initiated. In fact, there are indications that it might be better not to include a rationale for the behavior in the "then" part of implementation intentions, at least for easy tasks. Wieber, Sezer, and Gollwitzer (2014) compared the relative effectiveness of implementation intentions and goal intentions with regard to participants' mindset during a simultaneous tracking task and a go/no-go task (dual-task paradigm). First, Wieber and colleagues induced either a "why"- or a "how"-mindset in participants by means of an unrelated task. Afterwards, participants either formed a goal intention or internalized an implementation intention to perform well in the two simultaneous experimental tasks. Finally, the task difficulty was manipulated. For easy tasks, they found that participants in the why-mindset performed better when they had goal intentions than when they had internalized an implementation intention; vice versa, participants in the how-mindset performed better with implementation intentions than with goal intentions. For difficult tasks, the resuls were less clear: Mindsets became irrelevant for participants with goal intentions, whereas they were still important to participants with implementation intentions. Here, implementation intentions worked better with a how-mindset than with a why-mindset. Nevertheless, in general, implementation intentions were more effective than goal intentions for difficult tasks.

A good way to truly delineate the difference between implementation intentions and instructional prompts is to investigate the effectiveness of both types of support under varying levels of cognitive load. As mentioned above, the use of instructional prompts might actually constitute an additional cognitive load for learners (Bartholomé & Bromme, 2009; Horz et al., 2009). At the same time, implementation intentions have been shown to work even under conditions of high cognitive load (Brandstätter et al., 2001). Therefore, implementation intentions might serve as a favorable substitute for instructional prompts, especially under constraining conditions that make the initiation of effective multimedia processes difficult.

2.5 Research questions

Building on the foundation of these theoretical considerations, this dissertation addresses four overarching research questions by means of four experiments:

1) Can implementation intentions support effective cognitive processing in multimedia learning and thereby improve learning? Since implementation intentions have already been shown to be effective in many other contexts (cf. Gollwitzer & Sheeran, 2006), it is assumed that implementation can also support effective multimedia processes and thus improve learning. This research question stands at the core of this dissertation and is investigated in all four experiments.

2) In what way should implementation intentions be used to be most effective for multimedia learning? That is, which types of multimedia processes (i.e., text processing, picture processing, or integration) should implementation intentions support? How many implementation intentions should be used concurrently? Since the multimedia effect depends on learners integrating information across representational formats (Mayer, 2005; Schnotz, 2005), implementation intentions pertaining to integration processes are expected to be more effective than implementation intentions supporting text or picture processing. Moreover, a combination of implementation intentions pertaining to text processes, picture processes and integration processes are assumed to be the most effective. It has been shown that a concurrent use of four implementation intentions can work (Achtziger et al, 2008; De Vet et al., 2011); since more implementation intentions might lead to a more flexible use of effective cognitive processes, it is hypothesized that more implementation intentions would generally be better than only a single implementation intention. This research question was addressed mainly in Experiment 2 (see Section 4).

3) How do implementation intentions compare with other means to achieve similar goals, more specifically, to instructional prompts? Research has shown mixed results for instructional prompts (e.g., Horz et al., 2009; Kombartzky et al., 20010; Thillmann et al., 2009); one reason for this could be that instructional prompts may put additional cognitive demands on learners' working memory. Implementation intentions work regardless of learners' cognitive load however. Thus, it is expected that implementation intentions will compare especially favorably with instructional prompts under conditions of high cognitive load. Experiment 4 (see Section 6) was designed to answer this research question.

4) What is the relationship between implementation intentions and learners' motivation? Based on the compensatory effect of implementation intentions whenever volitional control is difficult, it is expected that implementation intentions will help especially those learners who

have little motivation to learn. This research question was focused on in Experiment 1 (see Section 3) but followed up on in the rest of the experiments.

In Sections 3 to 6, the four experiments of this dissertation will be described. Experiment 1 investigated the effect of implementation intentions on multimedia learning as well as on the interaction of implementation intentions with learners' task-specific motivation. Experiment 2 continued to address the effect of implementation intentions on multimedia learning by supporting the underlying cognitive processes. At the same time, the number of implementation intentions and the type of multimedia learning processes evoked by implementation intentions were varied. Experiment 3 was a replication of the main finding in Experiment 2 against a more conservative control condition. Finally, Experiment 4 aimed at delineating the differences between the use of implementation intentions and of instructional prompts by studying the effect of both under different conditions of cognitive load.

3 Experiment 1

In accordance with the Cognitive-Motivational Process Model, learners were expected to use effective multimedia processes more frequently when they had a high task-specific learning motivation, leading to better learning (Hypothesis 1; Berger & Karabenick, 2011; Vollmeyer & Rheinberg, 2006). Specifically, learners who think that they have a high probability of success, who have a high interest in the task, and who perceive the learning task as challenging, should show better learning outcomes. For anxiety, an inversed pattern was assumed in this experiment, so that learners who feel more anxiety would perform worse, since more fear of failure would most likely lead to less effective processing and thus to less learning (Bartels & Magun-Jackson, 2009).

In addition, it was hypothesized that implementation intentions would be effective in supporting the use of effective multimedia processing. Those learners who have internalized implementation intentions about the use of effective cognitive processes should show more frequent use thereof and consequently better learning outcomes than learners who have not (Hypothesis 2; Gollwitzer & Sheeran, 2006).

However, since implementation intentions have been shown to be effective especially when volitional control is difficult, it was further hypothesized that task-specific learning motivation plays an important moderating role when learning with implementation intentions (Hypothesis 3). More specifically, implementation intentions should especially help those learners who think that they have a lower probability of success, who are less interested in the learning task, and who feel less challenged by the task at hand, or who feel more anxiety and thus do not tend to use effective multimedia processes on their own.

3.1 Method

3.1.1 Participants and Design

Sixty students from the University of Tübingen participated in this study (mean age = 23.72 years, *SD* = 3.79; 44 female). In order to exclude participants with too much prior knowledge about this study's learning domain (cell division), students of biology and medicine were excluded from participating a priori. Participants received either course credits or 12 Euros as reimbursement for their voluntary participation.

Participants were randomly assigned to one of two experimental conditions. The study used a two-group experimental design (experimental and a control condition) with four motivational variables (probability of success, interest, challenge, and anxiety) as potential moderators.

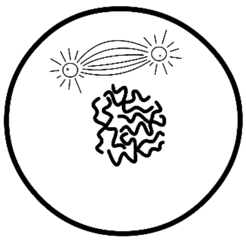
3.1.2 Materials

3.1.2.1 Learning materials

The printed learning materials consisted of a written and illustrated explanatory text about the topic of cell division, divided into three parts. First, the text gave a general and concise explanation on the role of cell division for life, on important cell structures, on DNA and its storage in the form of chromosomes, as well as on the specific features of gametes. Subsequently, the five phases of mitosis (interphase, prophase, prometaphase, metaphase, anaphase, and telophase) were explained, which result in the production of two genetically identical daughter cells, followed by an explanation of the eight phases of meiosis (prophase I, prometaphase I, anaphase I, telophase I, prophase II, metaphase II, anaphase II, and telophase II), which ends with the creation of four gametes; one ovum and three polar bodies in women, four spermatozoa in men.

Experiment 1

Each DIN A4 page of the learning material detailed only one topic or one phase of cell division and consisted of a picture at the top and the corresponding text below. The text had an overall length of 2,119 words. In the text, there were a total of 19 schematic color and black-and-white illustrations depicting the contents of the text, of which 12 contained short written labels (cf. Scheiter, Gerjets, Huk, Imhof, & Kammerer, 2009; Schüler, Scheiter, & Gerjets, 2013). Figure 2 shows an example from the learning materials.



Prophase

Die Mitose beginnt mit der Prophase. In der Prophase sehen die Zwei–Chromatid-Chromosomen im Zellkern zunächst knäuelförmig aus. Sie verdichten und verkürzen sich dann, bis sie schließlich im Mikroskop als Zwei-Chromatid-Chromosomen sichtbar werden.

Im Zellkern lösen sich die Nukleolen auf. Die Membran, die den Zellkern umgibt, beginnt in kleine Teile zu zerfallen.

Spätestens in der Prophase verdoppeln sich auch die Zentriole. Dies sind kleine körnchenartige Strukturen, die meist in der Nähe des Zellkerns liegen.

Die beiden Zentriole weichen auseinander und wandern zu den jeweils gegenüberliegenden Enden der Zelle, den Zellpolen. Zwischen den Zentriolen entsteht der Spindelapparat, der aus einzelnen Spindelfasern besteht. Die Zentriole bestimmen die Anordnung der Spindelfasern.

Figure 2. Example from the learning materials in Experiment 1.

3.1.3 Measures

Dependent variables encompassed participants' learning outcomes, both recall and transfer, while task-specific learning motivation (probability of success, interest, challenge, and anxiety) was measured as a moderator. Prior knowledge was measured as a control variable.

3.1.3.1 Motivation

Participants' motivation was assessed as a moderator by means of the Questionnaire on Current Motivation (QCM; Vollmeyer & Rheinberg, 2006). This questionnaire measures the current achievement motivation with regard to a specific task and has a total of 18 items distributed across four scales: probability of success (e.g., "I think I am up to the difficulty of this task."), interest (e.g., "For tasks like this I don't need a reward, they are lots of fun anyhow."), challenge (e.g., "I'm really going to try as hard as I can on this task."), and anxiety (e.g., "I feel under pressure to do this task well."). Each item has to be answered on a 7-point Likert-type scale (1 = "disagree" to 7 = "agree"). The QCM had an internal consistency of Cronbach's $\alpha = .70$ for probability of success, $\alpha = .70$ for interest, $\alpha = .66$ for challenge, and $\alpha = .85$ for anxiety.

In both conditions, the QCM was administered right before the learning phase; this was done to measure current task-specific learning motivation as closely to the actual learning task as possible and because it was not expected that the experimental manipulation would have an impact on learners' current motivation per se. The instruction for the QCM was slightly adapted from the original and read as follows (translated from German):

"Before the learning phase, we want to know your current attitude towards the task that was just described (i.e., learning the following contents and then taking a test). On this page, you will find a number of statements. For each, please mark the statement that best describes you."

3.1.3.2 Prior knowledge

The test of prior knowledge comprised three measures: First, two 4-point Likert-type scale items measured self-reported prior knowledge regarding mitosis and meiosis ("How much prior knowledge do you have concerning mitosis/meiosis?" 1 = "none" to 4 = "much"; cf. Mayer & Moreno, 1998). The two items were averaged for the purpose of data analysis.

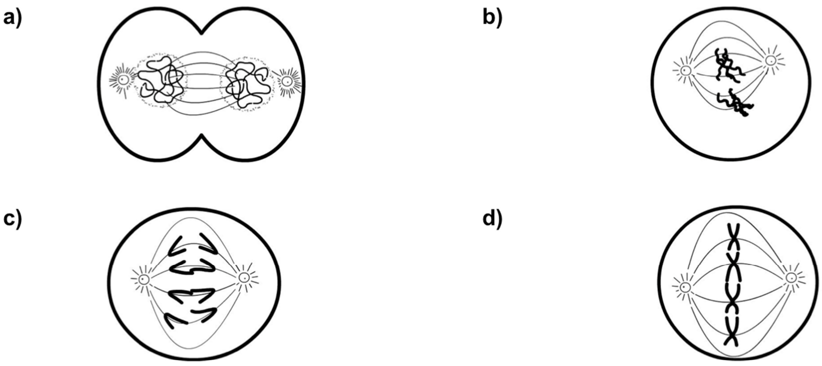
Second, participants were asked for the final grade in their most recent biology course (in accordance with the German grading system, 1 = "very good" to 6 = "unsatisfactory"). Third, participants' general prior knowledge in the life sciences was measured by 24 items from the Life Sciences scale in the Test of Basic Scientific Literacy (Laugksch & Spargo, 1996). These items are phrased as statements about scientific processes or interrelationships between scientific concepts (e.g., "The chemical processes in a cell are controlled both inside as well as outside the cell." or "Living organisms do not follow the same principles of conversation of matter and energy as other natural systems."). Learners have to rate these statements as either correct, incorrect, or "unknown" (meaning "I do not know the answer"). One point was assigned to each correct response, whereas no points were assigned to either an incorrect response or an "unknown" rating. The sum of assigned points was transformed into a percentage.

3.1.3.3 Post-test

The post-test measuring learning outcomes consisted of 32 multiple-choice items. Each of the items had four answer options, with one correct answer. Of the 32 items, 16 items assessed recall performance by requiring knowledge that had been explicitly stated in the learning materials, while the other 16 items tested transfer performance by requiring inferences regarding information that had not been explicitly contained in the learning materials. Furthermore, items varied in their presentation format in that they either consisted of a verbal question with verbal answers, a pictorial question with pictorial answers, a verbal question with pictorial answers, or a pictorial question with verbal answers. Figure 3 shows two examples from the post-test. One point was assigned to each correct response and the sum of correct responses was transformed to a percentage. The 16 recall items had an internal consistency of Cronbach's $\alpha = .50$, whereas the 16 transfer item had an internal consistency of $\alpha = .29$.

Recall

Which picture represents the last phase of mitosis?



Transfer

Assume that gametes are also formed by mitosis. Which consequences would this have on the process of reproduction?
a) The number of chromosomes would double from generation to generation.
b) Female meiosis would result in four gametes instead of only one gamete.
c) Chromosomes would not be split into chromatids.
d) There would be no consequences.

Figure 3. Examples of Recall and Transfer items in Experiment 1.

3.1.3.4 Use of effective multimedia processes

As a self-report measure for the use of the effective multimedia processes investigated in this study, two items asked specifically for the processes that were supported in this study ("In the learning phase, after having opened a new page, how often did you look thoroughly at the picture first?" and "In the learning phase, after having read a sentence, how often did you search the picture for the contents described therein?"). Answers had to be given on 7-point Likert-type scale (1 = "never" to 7 = "very often"). For the purpose of data analysis, the mean of both items was used.

Although self-report measures have been shown to not always result in accurate measurement of strategic cognitive processing (cf. Schellings & Van Hout-Wolters, 2011; Veenman, 2011), this type of measure was chosen over online measures like think-aloud protocols or eye-tracking because resources for conducting this study were limited; for instance, due to time and space constraints, four to six participants took part in this study in parallel, preventing the use of think-aloud protocols. Still, in the spirit of a multi-method approach, I attempted to measure learners' general use of strategic processing: learners were explicitly allowed, but not prompted, to make notes (highlights, underlining, comments, etc.) in the learning materials. The degree to which learners made notes was coded (partly by two independent raters) and analyzed but did not reveal any significant findings at all. In general, learners made only very few notes. Since this measure did not help in explaining any of the results of this study, it was excluded from the following analyses.

3.1.4 Procedure

Participants were tested individually, with a single session lasting about 1 hour. Four to six participants worked in parallel in the same room; however, no interaction was allowed between participants. It began with a short standardized introduction given by the experimenter. Afterwards, participants were handed a booklet of questionnaires for assessing demographic data, prior knowledge, and current task-specific learning motivation. In addition, the booklet contained a concise description of the experiment's structure and domain. Regardless of condition, all participants received the following instruction before the learning phase (translated from German):

"Cell division is a process that impacts us all on a daily basis. Without cell division, none of us could exist. Furthermore, erroneous processes during cell division can have grave consequences. Thus, it is important to know about how cell division works.

In order to understand the process of cell division correctly, you should make optimal use of the learning materials in the learning phase! Therefore, look at the pictures thoroughly and try to connect the text contents with the picture contents!"

The purpose of the first paragraph was to convey a sense of personal relevance to the participants, thus potentially increasing their engagement during the learning task (cf. Brophy, 1999). The second paragraph was included to ensure that participants in both experimental conditions had the same information about effective multimedia processes, thereby excluding the possibility that the knowledge of useful multimedia processes would confound any possible effects of implementation intentions.

In the experimental condition, participants were additionally introduced to the concept of implementation intentions and were asked to copy by handwriting two pre-phrased implementation intentions about the use of effective multimedia processes at that point. This instruction read as follows (translated from German):

“Implementation intentions are specific ‘if-then’ plans in which you connect a condition, under which you want to realize an action, with this action. *Example: ‘If I return home after work on Friday, then I will take my gym bag and will go to the gym!’* Implementation intentions can support you in making optimal use of the learning materials! Therefore, please write down both of the following implementation intentions five times each and resolve to realize the implementation intentions during the learning phase! It is important that you really want to realize the implementation intentions!”

Participants were instructed to write down the following two pre-phrased implementation intentions five times each (translated from German): *“If I have turned a page, then I will thoroughly look at the picture first.”* and *“If I have read a sentence, then I will search the picture for the contents described therein.”* This method of inducing the internalization of implementation intentions was based on Achtziger and colleagues (2008). In the phrasing of the implementation intentions, the “if-then” structure was deliberately chosen over a more naturalistic phrasing, as previous research had shown this type of phrasing to be more effective (Chapman, Armitage, & Norman, 2009).

It is important to note that, while there was no instruction to use particular multimedia processes in the control group, both groups received identical information with regard to the usefulness of thoroughly looking at a picture on each page first and of searching the picture for the information described in the text.

After they had finished and handed back the first booklet, participants were handed a second booklet consisting of the learning materials. In order to create a more naturalistic learning situation, participants were allowed to go back and forth through the pages and had no time limit for learning. Finally, participants were given a third booklet after they had handed back the second one; this final booklet encompassed the post-test and the questionnaire about the use of effective multimedia processes. Participants were allowed to work at their own pace. After the experiment, subjects were debriefed and given their respective remuneration.

Table 1: Means and standard deviations for learning outcomes, the four scales of the QCM as well as other control variables in Experiment 1.

	Control		**Implementation Intentions**	
	M	***SD***	***M***	***SD***
Recall performance (% correct)	66.46	16.53	71.04	11.55
Transfer performance (% correct)	46.25	11.79	46.04	11.89
Probability of success (1-7)	4.65	1.05	4.66	0.98
Interest (1-7)	3.53	1.11	3.99	0.92
Challenge (1-7)	4.62	0.97	4.92	1.08
Anxiety (1-7)	3.50	1.25	3.65	1.65
Self-reported prior knowledge (mitosis/meiosis; 1-4)	1.95	0.77	2.13	0.64
Last grade in Biology (1-6)*	2.40	1.16	2.00	0.79
Basic Scientific Literacy (Life Sciences; % correct)	74.86	12.01	77.36	11.72
Time-on-task (minutes) for the learning phase	26.23	8.20	30.63	10.85

* Note: In the German grading system 1 means „very good“, while 6 means „unsatisfactory“.

3.2 Results

3.2.1 Control variables

Univariate analyses of variance (ANOVAs) were conducted to test whether the two experimental groups differed in prior knowledge (see Table 1 for means and standard deviations of the control variables). The experimental groups did not differ in their self-reported prior knowledge in mitosis and meiosis, $F(1,58) = 1.00$, $MSE = 0.50$, $p = .32$, $\eta_p^2 = .02$, their last grade in Biology, $F(1,58) = 2.43$, $MSE = 2.40$, $p = .12$, $\eta_p^2 = .04$, or their score on the Life Sciences scale of the Test of Basic Scientific Literacy, $F < 1$. Finally, although the experimental group tended to spend more time on the learning phase, there was no reliable group difference in time-on-task in the learning phase, $F(1,58) = 3.14$, $MSE = 290.40$, $p = .08$, $\eta_p^2 = .05$. Table 1 lists the means and standard deviations of the time that learners spent on the learning phase.

3.2.2 Motivational variables

Since the QCM was administered after the experimental manipulation, a multivariate analysis of variance (MANOVA) with the four scales of the QCM as dependent variables was conducted to see whether there were differences between both experimental groups. There was no significant effect of experimental condition on the different scales of the QCM, Pillai's Trace = .07, $F(4,55) = 1.04$, $p = .40$. Means and standard deviations for the different scales can be found in Table 1.

3.2.3 Learning outcomes

Multiple regression analyses were conducted to analyze recall and transfer performance. The analyses each used nine predictors, entered simultaneously: the experimental condition (the control condition was coded as -1, the implementation intention condition was coded as +1), each of the four scales of the QCM (centered), and the interaction of experimental condition and each respective scale. Table 1 lists means and standard deviations of the learning outcomes.

Concerning recall performance, the overall regression model was significant, albeit only marginally, adj. $R^2 = .14$, $F(9,50) = 2.04$, $p = .054$. Table 2 lists all *B*-, *SE*-, β-, and *p*-values of this analysis. There was a marginal effect of implementation intentions, implying a trend of implementation intentions increasing learners' recall performance; furthermore, a significant interaction between implementation intentions and the scale interest emerged. The analysis showed no other significant effects. In order to interpret the interaction between implementation intentions and the scale interest, simple slopes analyses for -1 and +1 standard deviation of the continuous moderator interest were conducted (cf. Aiken & West, 1991). For learners with low scores on interest, a significant positive slope indicated that implementation intentions improved recall performance, $B = 8.03$, $SE = 2.60$, $\beta = .59$, $p = .002$, whereas no effect of implementation intentions was found for learners with a high interest score, $B = -3.00$, $SE = 2.49$, $\beta = -.21$, $p = .23$ (see Figure 4).

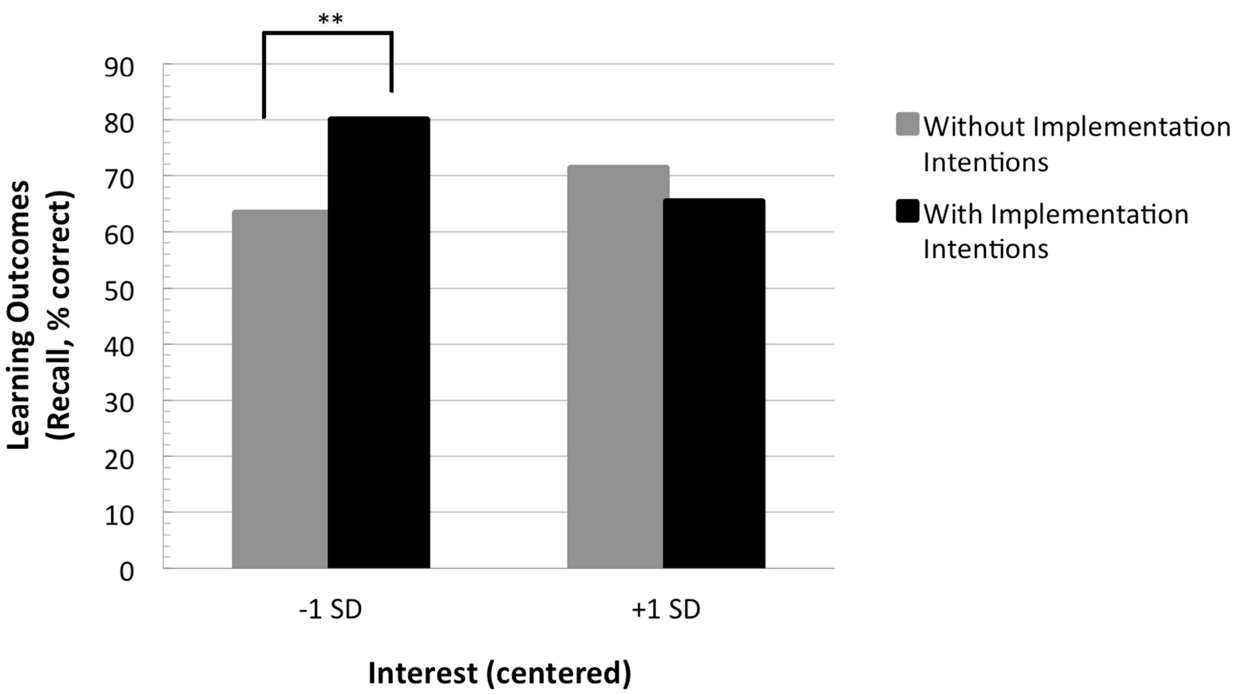


Figure 4. Recall performance (in percent correct); ** $p < .01$.

For transfer performance, the regression model was not significant, adj. $R^2 = .05$, $F(9,50) = 1.37$, $p = .23$, meaning that the model could not explain a sufficient portion of the variance in the data. As can be seen in Table 2, only interest showed a significant effect on transfer performance; surprisingly, learners with higher scores on the scale interest were less successful in answering transfer questions than learners with low scores on interest. Furthermore, there was a marginally significant effect for the scale probability of success, in that learners with higher scores on probability of success tended to perform better on the transfer items.

3.2.4 Use of effective multimedia processes

In order to investigate whether implementation intentions had an effect on the self-reported use of effective multimedia processes, a multiple regression analysis was conducted in order to investigate differences between the two experimental conditions, with experimental condition (coding as reported above), each of the four centered scales of the

QCM, as well as the four interaction terms (experimental condition × scales) as simultaneous predictors. The regression model was significant, adj. R^2 = .19, F(9,50) = 2.56, p = .02 (see Table 2 for all *B*-, *SE*-, β-, and *p*-values). The group with implementation intentions (*M* = 5.95, *SD* = 1.22) reported having used the effective multimedia processes in question more frequently than the control group (*M* = 5.05, *SD* = 1.35). With regard to the scales of the QCM, there was a significant effect of interest on the self-reported use of effective multimedia processes in that more interested learners reported having used these processes less often. At the same time, a significant effect of challenge indicated that more challenged learners reported higher use of the multimedia processes in question. The analysis showed a positive trend on the scale probability of success in that learners with higher scores in probability of success tended to report a more frequent use of effective multimedia processes than learners with lower scores on that scale. Moreover, there was no significant interaction between experimental condition and the scale interest.

That said, self-reported use of effective multimedia processes correlated neither with recall performance (*r* = .07, *p* = .58) nor with transfer performance (*r* = .03; *p* = .82).

Table 2: B-, *SE*-, β-, and *p*-values of all nine predictors, for recall performance, transfer performance, and mean self-reported strategy use in Experiment 1.

		B	***SE***	**β**	***p***
Recall performance	Experimental condition	3.01	1.79	0.21	0.10
	Probability of Success	-1.22	2.10	-0.09	0.56
	Interest	-1.59	1.99	-0.12	0.43
	Challenge	-1.66	1.87	-0.12	0.38
	Anxiety	-1.30	1.38	-0.13	0.35
	Experimental condition × Probability of Success	-3.35	2.10	-0.24	0.12
	Experimental condition × Interest	-5.04	1.99	-0.36	0.02
	Experimental condition × Challenge	1.25	1.87	0.09	0.51
	Experimental condition × Anxiety	1.06	1.38	0.11	0.45
Transfer performance	Experimental condition	0.71	1.53	0.06	0.64
	Probability of Success	3.42	1.80	0.29	0.06
	Interest	-3.81	1.71	-0.34	0.03
	Challenge	-0.003	1.61	0.00	1.00
	Anxiety	0.62	1.18	0.08	0.60
	Experimental condition × Probability of Success	0.50	1.80	0.04	0.78
	Experimental condition × Interest	-1.75	1.71	-0.15	0.31
	Experimental condition × Challenge	-1.05	1.61	-0.09	0.52
	Experimental condition × Anxiety	1.49	1.18	0.18	0.21
Self-reported strategy use	Experimental condition	0.49	0.15	0.41	0.002
	Probability of Success	0.33	0.17	0.28	0.06
	Interest	-0.38	0.16	-0.32	0.02
	Challenge	0.31	0.15	0.26	0.05
	Anxiety	0.003	0.11	0.004	0.98
	Experimental condition × Probability of Success	-0.19	0.17	-0.16	0.28
	Experimental condition × Interest	0.28	0.16	0.23	0.10
	Experimental condition × Challenge	-0.12	0.15	-0.10	0.44
	Experimental condition × Anxiety	-0.01	0.11	-0.01	0.96

3.3 Discussion

The combined presentation of text and pictures can only lead to a better understanding if learners actively process and integrate the information of both types of representation. However, self-regulated, active processing of multimedia materials is demanding and thus needs to be sufficiently motivated, so that learners may often fail to make best use of these materials. This study investigated the question of whether implementation intentions would improve learning from multimedia learning materials by fostering the use of effective cognitive processes and whether this effect would be moderated by learners' achievement motivation.

Contrary to Hypothesis 1, none of the four motivational factors probability of success, interest, challenge, or anxiety had a main effect on recall performance. The averaged means and standard deviations of the QCM scales indicate that learners were moderately motivated regarding this learning task, so these results can hardly be attributed to a general lack of task-specific learning motivation. Note, however, that all scales displayed a comparatively small statistical variation (cf. Rheinberg, Vollmeyer, & Burns, 2001), which may have made regression analyses more difficult. A possible explanation is that the measurement of initial motivation, as assessed by the QCM, is not sufficient to show a statistical impact on learning outcomes. Vollmeyer and Rheinberg (2006) found in their study that initial motivation's impact on performance was mediated by participants' motivational state, as assessed during rather than prior to learning, which (together with strategic processing) influenced learners' knowledge acquisition. If learners' initial motivation did not persist and/or transform into an increased motivational state during the complex learning task in this study, it could explain why there was not a measurable effect on recall performance. Since task-specific learning motivation was only measured once in this study, this explanation can neither be confirmed nor denied at this point.

For transfer performance, on the other hand, no comparable pattern of results was found. Here, anxiety and challenge had no significant impact on performance, whereas there was a marginally significant positive effect of probability of success on learners' transfer performance, thereby at least partly confirming Hypothesis 1. Interest, on the other hand, even had a contradictory effect: an increased interest led to worse transfer performance. This paradoxical effect is quite puzzling, and I currently lack a satisfying explanation for it; it remains to be seen whether this effect can be replicated in future studies.

Concerning Hypothesis 2, there was only a marginal trend that implementation intentions improved learners' recall performance. Then again, implementation intentions did not have a main effect on transfer performance. This raises the question why implementation intentions did not have an effect on transfer performance at all. At this point, it should be noted that the internal consistency of the 16 transfer items was very low; thus, any lack of effects could, at least partly, be explained by the unreliable testing method for transfer performance. As a direct consequence, different items and measures were used in subsequent experiments (Experiments 2 to 4).

However, there might also be a theoretical reason why there was no effect of implementation intentions on transfer performance. At least one of the two effective multimedia processes in the pre-phrased implementation intentions aimed at supporting the integration of text and pictures. Therefore, implementation intentions should have led to the construction of a more comprehensive mental model, resulting in more inferences and better transfer performance. At the same time, the results for recall show that implementation intentions did have some effect on learning, at least for less interested learners. Taken together, implementation intentions supported recall under specific circumstances, yet they still failed to unfold their full potential for facilitating text-picture integration despite addressing a specific integration process. A possible explanation of this result could be that just evoking the integration of text and picture elements is not enough. According to the theories of multimedia learning, such as the CTML or the ITPC, the integration of information from verbal and pictorial mental models with prior knowledge is a higher-order process of understanding multimedia materials, preceded by a number of representationally specific processing steps. Thus, in order for this active integration process to succeed, both the information in the text and the picture must have been processed on a lower level before. Accordingly, maybe it is not sufficient to only support this last step in mental model construction, but possibly it is necessary to equally support the cognitive processes preceding integration as well. Based on these considerations, Experiment 2 investigated whether an approach that supports a wider spectrum of effective multimedia processes would lead to the construction of a more comprehensive mental model and thus to better performance.

With respect to the interaction hypothesis (Hypothesis 3), there were different results for recall and transfer performance. For recall performance, only one of the expected interactions between implementation intentions and the four motivational factors emerged: especially those learners with low interest in the task significantly benefited from the use of implementation intentions. It makes sense that this particular interaction proved to be significant, considering that value (or interest) has been shown to be an important predictor for the use of strategic processing over a longer span of time (Pokay & Blumenfeld, 1990). Thus, the interaction was in accordance with Hypothesis 3, suggesting that less interested learners would have more difficulty with using effective multimedia processes than more strongly interested learners and that implementation intentions would therefore prove helpful particularly for these learners. Hypothesis 3 was based on the findings that implementation intentions are especially effective under volitionally difficult circumstances by acting as a substitute for the volitional control that would result from being motivated to act (e.g., Gollwitzer & Brandstätter, 1997). So why did implementation intentions not interact with the other three scales of the QCM? One distinction that might be drawn between the scale interest and the other three scales of the QCM is the degree to which they depend on the specific task vis-à-vis the learning domain. Learners' interest in the learning task is much more likely to be influenced by their interest in the topic (i.e., the learning domain), while the perceived probability of success, level of challenge, and performance anxiety will more likely be dependent on the task itself. At the same time, participants were aware of the experimental situation and knew that there would be no negative consequences regardless of how well they performed. It might be possible that implementation intentions only interacted with interest because this scale measured a deeper and more personal engagement with the underlying domain, while the specific learning task, with its associated difficulty and challenges, was considered less meaningful in the grand scheme of things. For transfer performance, there were no significant interactions between implementation intentions and the four motivational factors.

In summary, whereas implementation intentions did not improve learning outcomes in general, they seemed to be suited to get less interested learners on board for the learning task, thus helping especially those learners who saw the least value in the learning task. This finding is promising and interesting insofar as the intervention of phrasing and internalizing an implementation intention is extremely simple and can be very easily implemented in ecologically valid settings such as the classroom, as other research on implementation intentions has shown (Gollwitzer & Sheeran, 2006).

In the context of multimedia, these first results seem promising since it might be an cost-effective way to somewhat counteract the commonly found "Matthew effect", that is the finding that, without compensative instructional measures, those learners with favorable learner characteristics such as higher motivation or better spatial abilities often profit more from text-picture combinations than those with less favorable learner characteristics (e.g., Höffler, 2010). In this sense, implementation intentions could be used to "even out the odds" a bit at very little cost.

Two questions remain, however. Why did implementation intentions not work for learners with higher interest? Why did the self-reported use of effective multimedia processing not correlate with learning outcomes?

The fact that less interested learners profited more from implementation intentions than those with higher interest was expected. However, my expectation that more interested learners would also benefit from implementation intentions albeit to a lesser degree did not hold. One explanation for this result is that learners who are already involved in the learning task, may have used habitual cognitive processes regarding multimedia learning on their own. Then, the triggering of cognitive processes via implementation intentions might have conflicted with learners' more habitual processes. Another explanation for this result might lie in the way learners were instructed to internalize the two pre-phrased implementation intentions. More specifically, having to write down the same implementation intention five times each might have resulted in reactance, especially for those learners who were already interested in the task. Differently phrased, the instruction might have been too intrusive resulting in an undermining of the implementation intention effect for more interested learners. Consequently, the implementation intention instructions were slightly changed in subsequent studies in order to avoid possible negative effects (Experiment 2 to 4).

Second, the question why self-reported use of effective multimedia processes did not correlate with learning outcomes might be explained by the measure used to assess the use of these processes. In Experiment 1, a retrospective self-report measure was used. One explanation is certainly that by internalizing the implementation intentions, the experimental group became more sensitized to their use of these processes and thus tended to overestimate their use of them. Thus, the use of effective multimedia processes as measured by self-reports would not reflect actual behavior and therefore would not correlate with the learning outcomes. Another explanation is that such a measure necessitates that the information about the use of cognitive processes is consciously accessible for learners. It can be argued, however, that the use of effective cognitive processes (especially when automatized by the use of implementation intentions) is not necessarily conscious and thus not easily accessible for self-report (cf. Veenman, 2011). Therefore, this poses the question whether learners really had access to the information about their use of effective multimedia processes. In any case, future studies need more objective process measures that work even in case the use of effective multimedia processes should not be consciously accessible. For this purpose, eye-tracking or verbal protocols offer a more valid and reliable way of assessment (Scheiter & Van Gog, 2009; Van Gog & Scheiter, 2010). For this reason, eye-tracking was used both in Experiments 2 and 4 to gain insight into learners' cognitive processing during learning.

The present experiment showed that implementation intentions about the use of effective multimedia processes could support recall performance for learners with low interest in the task. At the same time, implementation intentions did not have an effect on transfer performance. This indicates that learners might need a broader support in order to build an integrated, coherent mental model. Therefore, the next step was to find out whether a combination of alternative effective multimedia processes would be more effective for learning when supported by implementation intentions and which combination would result in the best learning outcomes. The choice of the two implementation intentions used in this study (and the multimedia processes contained therein) was primarily based on theoretical considerations. However, these two implementation intentions did not cover the whole spectrum of effective multimedia processes, that is, text processing, picture processing, and text-picture integration. On this ground, comparing the effectiveness of implementation intentions with all three types of processes, as well as a combination of them seemed like a necessary step. Another question is intrinsically intertwined with the previous question: how many implementation intentions can and should be used concurrently? Prior research has shown that more than one implementation intention can be used concurrently (e.g., Achtziger et al., 2008; De Vet et al., 2011), but the optimal number of implementation intentions is still an open empirical question. Moreover, a combination of text, picture, and integration processes would require learners to internalize at least three implementation intentions concurrently. Accordingly, Experiment 2 investigated the effectiveness of different types of effective multimedia processes when evoked by implementation intentions as well as the effectiveness of different numbers of concurrently used implementation intentions.

In conclusion, implementation intentions seem to represent a promising means to support the use of effective multimedia processes, especially for learners with less suitable learner characteristics, such as low task interest. They are very easy to adapt to and implement in a wide variety of educational contexts, making them an attractive self-regulation technique in volitionally difficult situations.

4 Experiment 2[1]

In many regards, Experiment 1 was a first attempt to investigate whether implementation intentions would facilitate the use of cognitive processing during multimedia learning. Although it gave first indications that implementation intentions can improve learning outcomes, the results were not fully conclusive, either with regard to how much learning outcomes were affected or how implementation intentions interacted with learners' achievement motivation. Hence, the purpose of Experiment 2 was to zoom in on the effect of implementation intentions on effective multimedia processing, while setting aside the moderating role of learners' achievement motivation for the time being[2]. Instead, Experiment 2 tried to answer some of the questions raised by the findings in Experiment 1, such as how implementation intentions should best be implemented for supporting an effective processing of multimedia materials, thereby achieving the maximum impact on learning outcomes. Since learners' achievement motivation proved itself as a potential moderator, however, it nevertheless was assessed and controlled for in Experiment 2 and the subsequent experiments.

The primary goal of Experiment 2 was to revisit my hypothesis that implementation intentions can successfully facilitate the initiation of effective cognitive processes during multimedia learning (Hypothesis 1).

---

[1] Results from Experiment 2 and 3 are reported in Stalbovs, Scheiter, and Gerjets (2015).

[2] Testing the moderating effect of motivation on the use of implementation intentions in this experiment would have presented some problems. Statistically, the increase of predictors would have required more participants per experimental group. Moreover, it would have been difficult to test motivation during the learning phase (cf. Experiment 1) due to the use of eye-tracking.

Experiment 2

Additionally, this experiment aimed at determining what type of multimedia-specific cognitive processes (i.e., text processing, picture processing, text-picture integration, or a combination thereof) should best be supported. Since a full understanding of a given multimedia content can only occur with the integration of information across representational formats (Mayer, 2005; Schnotz, 2005), I expected implementation intentions about integration processes to be more successful than those about text or picture processing (Hypothesis 2).

In Experiment 1, the implementation intentions focused only on picture processing and integration and had an impact, albeit weak, on recall performance. This finding suggested that learners might need implementation intentions that address both representational formats individually as well as their integration. Thus, I expected a combination of implementation intentions about three types of processes to be the most successful (Hypothesis 3).

Strongly intertwined with this issue is the question of how many implementation intentions can be used concurrently. Whereas most studies investigating implementation intentions used only a single implementation intention, the concurrent use of more than one implementation intention has been shown to work (e.g., Achtziger et al., 2008; De Vet et al., 2011). I expected that the use of more than one implementation intention would lead to a more flexible use of effective multimedia processes and, thus, to better learning outcomes (Hypothesis 4).

Since the use of self-report measures to assess the use of effective cognitive processes proved to be unsuited for the task in Experiment 1, the present experiment used eye-tracking to measure learners' attention allocation and sequence of cognitive processing (cf. Just & Carpenter, 1980). With regard to gaze data, I expected that the internalization of implementation intentions would have an effect on fixation times regarding text and pictures as well as on gaze transitions between corresponding text paragraphs and picture elements.

More specifically, I expected that, compared with the control group, there would be longer fixation times on text for learners who learned with implementation intentions pertaining to text processes, especially if they learned with more than one implementation intention. Those learners who learned with implementation intentions regarding integration processes (regardless of how many) should show increased fixation times on the text, since an active processing of text acts as prerequisite for integration. Moreover, learners who learned with implementation intentions pertaining to a combination of cognitive processes should show the longest fixation times on text, since they should be more actively engaged with the learning materials in general (Hypothesis 5).

In an analogous manner, compared with the control group, I expected longer fixation times on pictures for learners who learned with implementation intentions with regard to picture processes, especially if they learned with more than one implementation intention. As above, learners who learned with implementation intentions concerning integration should also show increased fixation times on pictures. The longest fixation times on pictures were expected for learners who learned with implementation intentions pertaining to a combination of cognitive processes, for the same reasons as noted above (Hypothesis 6).

With regard to gaze transitions between text and pictures, compared with the control group, I expected that there would be an increase for learners who internalized implementation intentions regarding integration processes, especially for those who learned with more than one implementation intention. Furthermore, I expected the group with implementation intentions regarding a combination of cognitive processes to look back and forth between text and pictures as often as the group with several implementation intentions regarding integration, since they should be more engaged with the learning material in general and should benefit from a widespread use of effective multimedia processes (Hypothesis 7).

Finally, I expected that these three gaze measures would predict learning outcomes (Hypothesis 8).

While learners' achievement motivation during learning was shown to partially moderate the effectiveness of implementation intentions in Experiment 1, this motivational influence was not a specific focus in Experiment 2. Nevertheless, learners' task-specific learning motivation was assessed as a control variable before learning and expected to have an impact on learning outcomes.

4.1 Method

4.1.1 Participants and Design

160 students from the University of Tübingen participated in this experiment (mean age = 23.21 years, *SD* = 3.49; 128 female). Like in Experiment 1, students of biology, medicine, or of any field with a heavy focus on biology were excluded a priori from participating since the experiment's learning domain pertained to the biological process of cell division. Participation was voluntary and was reimbursed with either 12 Euros or course credits. Participants were randomly assigned to one of eight experimental conditions with 20 students serving in each condition.

The experiment used a between-subjects 2×3 experimental design with two additional conditions. With regard to the first factor "number of implementation intentions", conditions differed in whether learners had to internalize one pre-phrased implementation intention (1) or three pre-phrased implementation intentions (3). With regard to the second factor "type of supported cognitive processes", three different types of implementation intention were compared: implementation intentions fostering text processes (TXT), picture processes (PIC), or text-picture integration processes (INT). In addition, there was a group that learned with three implementation intentions, each supporting a different type of multimedia learning process; students in this group learned with one implementation intention fostering a text process, one evoking a picture process, and a third eliciting a text-picture integration process (MIX). Finally, there was a control group (CTRL) in which students neither learned with any implementation intentions nor received any information about effective multimedia learning processes.

For ease of reading, these eight groups will be referenced in the following as 1-TXT, 1-PIC, 1-INT, 3-TXT, 3-PIC, 3-INT, MIX, and CTRL, respectively.

4.1.2 Materials

4.1.2.1 Learning materials

The learning materials were a revised version of the learning materials used in Experiment 1; they consisted of a written and illustrated explanatory text about the topic of cell division and were presented on a computer screen with a resolution of 1650×1050 pixels. The learning materials were distributed across 19 presentation slides, each of which detailed one topic or one phase of cell division. Each slide consisted of a picture on the left and the corresponding text on the right. The text had an overall length of 2,049 words, containing a total of 19 schematic color and black-and-white illustrations depicting the contents of the text. It was segmented into semantically meaningful paragraphs, whereby each paragraph referred only to a single fact or concept, so that whenever a new concept was introduced, also a new paragraph began. None of the paragraphs referred to more than a single element in the picture. An example from the learning materials can be found in Figure 5.

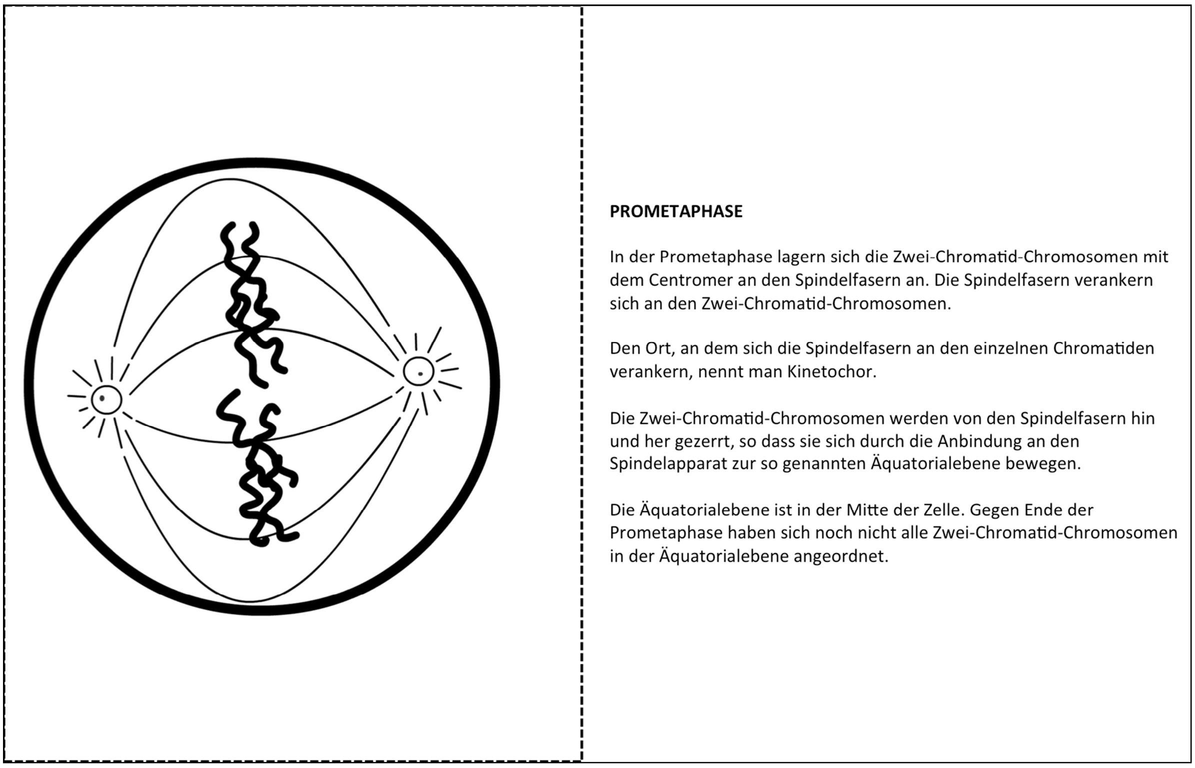


Figure 5. Example from the learning materials, explaining the prometaphase of mitosis.

4.1.3 Measures

Dependent variables for this experiment encompassed participants' learning outcomes and learners' gaze data; prior knowledge, task-specific learning motivation (probability of success, interest, challenge, and anxiety), and learning time served as control variables.

4.1.3.1 Motivation

Like in Experiment 1, participants' motivation was assessed as a control variable by means of the Questionnaire on Current Motivation (QCM; Vollmeyer & Rheinberg, 2006).

4.1.3.2 Prior knowledge

In order to get an accurate estimation of learners' prior knowledge, three measures were used. Two of these, learners' self-reported prior knowledge and scientific literacy in the life sciences, were the same measures used in Experiment 1 (see 3.1.3.2). Learners' self-reported prior knowledge regarding mitosis and meiosis was assessed by two 4-point Likert-type scale items (cf. Mayer & Moreno, 1998). For the purpose of data analysis, the two items were averaged. Learners' scientific literacy again was operationalized by 24 items from the Life Sciences scale of the Test of Basic Scientific Literacy (Laugksch & Spargo, 1996). The sum of assigned points was transformed into a percentage.

The third and newly introduced measure was a 13-items multiple-choice test on cell biology. Each of the items had four possible answers of which only one was correct (e.g., "What is a component of an animal cell? a) Chloroplasts, b) Ribosomes, c) Murein, d) Cellulose" or "Which component of a cell is dissolved almost completely during mitosis? a) Cell wall, b) Cell membrane, c) Nuclear membrane, d) Nucleotide"). One point was assigned to each correct response and the sum of correct responses was transformed into a percentage.

4.1.3.3 Post-test

The computer-based post-test measuring learning outcomes encompassed a total of 60 verification items in three categories (20 per category): items consisted either of a verbal statement (text items), a picture (picture items), or a picture combined with a verbal statement (integration items), which learners had to rate as true or false. See Figure 6 for an example of each category. The statements could be successfully verified from the text alone, the picture items could be successfully verified from the information in the pictures alone, whereas the combination of statement and picture could only be successfully verified by connecting the information from the text with the information from the pictures. Consequently, the first two categories (i.e., text and picture items) of the post-test assessed recall, whereas the third category (i.e., integration) measured comprehension of the multimedia learning content. That is, for a correct answer, the recall item required only a single fact mentioned in the text, displayed in the picture, or both. The comprehension items, however, necessitated learners to make connections between multiple facts in order to judge their validity. The 20 comprehension items of this post-test often, but not always, tested learners' understanding of commonalities and differences between mitosis and meiosis. The differentiation between mitosis and meiosis has been shown to be one of the major challenges in this domain for students at all levels (e.g., Flores, Tovar, & Gallegos, 2003; Lewis, Leach, & Wood-Robinson, 2000). Furthermore, for correctly answering the comprehension items, learners had to integrate a number of facts presented in the item first, that is, the correct answer required learners' reasoning, rather than recall. For instance, in order to correctly answer the integration item in Figure 6, learners had to know that chromosomes must be aligned along the equatorial plane before they can separate. (Otherwise, it would result in an unequal distribution of chromosomes between the two daughter cells.) At the same time, they need the pictorial information in the item to make that judgment.

One point was assigned to each correct response and the sum of correct responses was divided by the total number of items (mean accuracy). The post-test was presented via the E-Prime Professional software by Psychology Software Tools® (version 2.0).

Since this post-test was developed for this experiment and used for the first time, I tested for items that were either too difficult or systematically misunderstood by learners, as indicated by performance significantly below chance level. T-tests against chance level were conducted for each item in order to identify those items whose mean accuracy was significantly below chance level ($p < .05$). This was done both in Experiment 2 and in Experiment 3 (the first two experiments using this post-test). Those items whose mean accuracy was significantly below chance level both in Experiment 2 *and* in Experiment 3 were excluded from further analyses. Of the 60 items, 55 were answered at or above chance level (19 text items, 16 picture items, and 20 integration items). In Experiment 2, the post-test had an internal consistency of Cronbach's $\alpha = .68$. Given the satisfactory internal consistency of the items, I decided to collapse them into one performance measure in order to facilitate data analyses. Thus, verification accuracy reflects the ability to recall as well as comprehend information regarding mitosis and meiosis irrespective of item format (text, picture, or both).

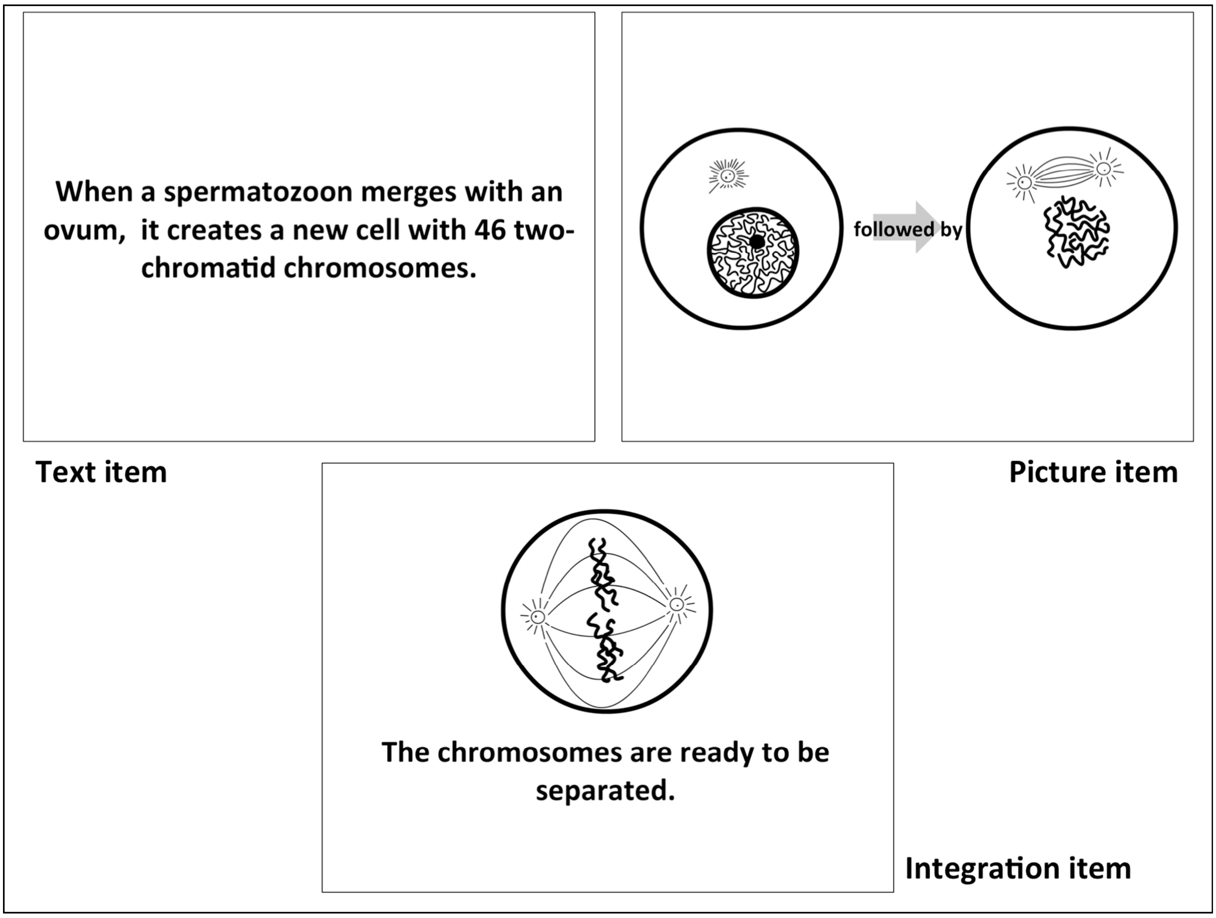


Figure 6. Examples from the post-test (translated from German).

4.1.3.4 Gaze data

Next to participants' learning outcomes, their eye movements constituted another dependent variable in this experiment. For the analysis of learners' eye movements, areas of interest (AOIs) were defined on every page of the learning materials. These AOIs encompassed the text as a whole and the picture as a whole. Furthermore, there were AOIs for every text paragraph (i.e., every idea unit) and every picture element that was referred to in the text; only those gaze transitions between text and picture were used that were between corresponding text paragraphs and picture elements (cf. corresponding transitions, Johnson & Mayer, 2012). Moreover, gaze transitions between text and picture always included the transitions from text to the picture and *vice versa*. For the purpose of data analysis, gaze transitions were added up across all pages of the learning materials.

In total, three eye movement measures were taken into consideration. Table 3 lists all measures, each measure's hypothesized association regarding multimedia processes, and the hypothesized differences on these measures between groups.

4.1.4 Apparatus

During the learning phase, learners' eye movements were recorded as process measures. For this, I used a SensoMotoric Instruments (SMI) RED250 remote eye-tracking system with a sampling frequency of 250 Hz. The eye-tracking system was controlled via the SMI iViewX™ software (version 2.5), while the experiment was presented via the SMI Experiment Center™ software (version 2.5). The system used a high-speed event detection algorithm with a peak velocity threshold of 40°/s and a minimum fixation duration of 50 ms.

Table 3: All three eye movement measures, their hypothesized correspondence to multimedia processes, and hypothesized group differences in Experiment 2.

Eye movement measure		**Hypothesized association with multimedia process**	**Hypothesized differences between groups**
Total fixation time	*All text paragraphs*	Amount of reading	CTRL, 1-PIC, 3-PIC < 1-TXT, 1-INT, 3-INT < 3-TXT < MIX
	All relevant picture elements	Amount of picture processing	CTRL, 1-TXT, 3-TXT < 1-PIC, 1-INT, 3-INT < 3-PIC < MIX
Transitions	*Between text paragraph and corresponding picture element, and vice versa*	Integrating the information in the text with the information in the picture, and vice versa	CTRL, 1-TXT, 1-PIC, 3-TXT, 3-PIC < 1-INT < 3-INT, MIX

4.1.5 Procedure

Participants were tested individually in a single session that lasted about 90 minutes. First, the experimenter gave participants a short standardized introduction. Afterwards, participants answered a computer-based questionnaire for assessing demographic data, their prior knowledge, and task-specific motivation. The computer-based questionnaire ended with a short description of the experiment's structure and domain.

In the experimental conditions, participants were then handed out a paper-and-pencil instruction that introduced the concept of implementation intentions. Depending on condition, this instruction differed with regard to how many implementation intentions were included and what type(s) of multimedia learning processes was supported by means of implementation intention. The instruction read as follows (translated from German):

"Have you ever experienced that you strongly intended to do something but did not do it in the end? This happens often to almost everybody! There is something, however, that can help you in these cases! Implementation intentions are very specific 'if-then' plans that are supposed to help you translate an intention (e.g., 'I want to go to the gym!') into an action. For having an implementation intention, you link a condition under which you want to perform an action with said action. Example: *'If I come home after work on Friday, then I take my gym bag and visit the gym!'* In the course of this experiment, please learn with the intention and the implementation intention to make optimal use of the learning materials! Implementation intentions can help you to make best use of the learning materials! Therefore, please copy the following implementation intention(s) twice and imagine how you will realize the implementation intention(s) later in the learning phase! It is important that you really want to realize the implementation intention!"

Experiment 2

At the end of the instruction, one or three pre-phrased implementation intentions were listed. Participants were asked to copy the listed pre-phrased implementation intentions twice by hand.

Compared with the implementation intention instruction in Experiment 1 (see Section 3.1.4), the instruction in this study gave an expanded explanation of implementation intentions. Moreover, learners were instructed to imagine themselves realizing the implementation intention(s). Finally, in order to minimize the risk of provoking reactance in learners, they were asked to write down each implementation intention only twice (instead of five times like in Experiment 1).

In total, nine different implementation intentions were used in this experiment; Table 4 gives an overview of all the pre-phrased implementation intentions. The MIX condition contained the following three implementation intentions:

1) Text process: *"If I have finished reading a page, then I will carefully re-read all paragraphs!"*
2) Picture process: *"If I am looking at a picture, then I will search for its central elements with regard to content!"*
3) Integration process: *"If I have read a paragraph, then I will search the picture for the contents described therein!"*

In the conditions with only one implementation intention, the selection of the singular implementation intention among the three of one type of multimedia process was counterbalanced between participants, in that the selection of the specific implementation intention was rotated between participants. Since participants had to imagine using the implementation intention(s), they were shown the first page of the learning materials for a short time in order to have an impression of how the learning phase would look like. After they had finished copying the implementation intentions, the eye-tracking system was calibrated for assessing participants' eye movements.

Experiment 2

In the control condition, participants did not receive the written implementation intention instruction or any further information regarding multimedia learning processes and went on to the eye-tracking calibration directly after the computer-based questionnaire. Following the calibration, the learning phase started. Participants were given no time limit during the learning phase and were not allowed to go back to previous pages. After the learning phase, participants answered the computer-based post-test. They had no time limit for the post-test but were instructed to work as quickly and accurately as possible. Finally, they were debriefed and given their respective remuneration.

Table 4: Overview over all nine multimedia processes and implementation intentions in Experiment 2.

Multimedia Processes	**Implementation Intention**
Text processes	
Overview / Global text processing	“If I have opened a new page, then I will carefully study the title first!”
Rehearsal	“If I have finished reading a page, then I will carefully re-read all paragraphs!” *
Organization	“If I have read a paragraph, then I will search for references to previous paragraphs!”
Picture processes	
Overview	“If I have opened a new page, then I will carefully study the picture first!”
Selection	“If I am looking at a picture, then I will search for its central elements with regard to content!” *
Organization	“If I have looked at a picture, then I will put its central elements into context with each other!”
Integration processes	
Integration text-picture	“If I have read a paragraph, then I will search the picture for the contents described therein!” *
Integration picture-text	“If I have looked at a picture, then I will search the text for explanations of the examined picture elements!“
Matching of mental models	“If I want to click to the next page, then I first carefully study the picture to verify my understanding of the text!”

The implementation intentions marked with an * were used in the mixed condition (MIX)

4.1.6 Data analyses

The data from the dependent variables were analyzed using effect coding (cf. Abelson & Prentice, 1997; Niedenthal, Brauer, Robin, & Innes-Ker, 2002). This was done for two reasons: first, I had specific, directional hypotheses about how the experimental manipulation would affect the learning outcomes and the gaze data. Second, effect coding allowed us to take the complex experimental design (2×3 design with 2 additional conditions) into account without having to break down the analysis into multiple smaller tests, which would have resulted in a loss of statistical power and alpha-error inflation.

The basic idea behind contrast coding is to test whether a specific model ("focal contrast"), which is based on the hypothesized relative group differences, better fits the observed data than a number of independent (i.e., orthogonal), alternative models ("residual contrasts") (Abelson & Prentice, 1997). If the focal contrast fits the data to a significant degree while the residual contrasts do not, it can be concluded that the hypothesized pattern of group differences describes the observed data accurately. If the focal contrast does not fit the data while the residual contrasts do, then the data do not conform to the hypotheses and are better explained by other models. If both the focal contrasts and the residual contrasts fit the model significantly, then the hypothesized group differences can be found in the data but the data are additionally explained by other patterns of relative group differences.

In effect coding, the relative differences of codes are meaningful. A coding of 0 represents the grand mean of the observed data, whereas codings of either below or above 0 represent relative deviations from the grand mean. Positive effect codes mean that the condition that has been assigned the code is expected to score above the grand mean, whereas a negative code means that the condition is expected to score below the grand mean.

Table 5: Effect coding of focal and residual contrasts for the analysis of learning outcomes in Experiment 2.

	1-TXT	**1-PIC**	**1-INT**	**3-TXT**	**3-PIC**	**3-INT**	**MIX**	**CTRL**
Focal contrast 1: Hypotheses 1 and 3	0	0	0	0	0	0	+1	-1
Focal contrast 2: Hypothesis 2	-1	-1	+2	-1	-1	+2	0	0
Foal contrast 3: Hypothesis 4	-1	-1	-1	+1	+1	+1	0	0
Residual contrast 1	-1	+1	0	-1	+1	0	0	0
Residual contrast 2	+1	-1	0	-1	+1	0	0	0
Residual contrast 3	-1	-1	-1	-1	-1	-1	+3	+3
Residual contrast 4	+1	+1	-2	-1	-1	+2	0	0

Experiment 2

Since there were eight different groups in this experiment, seven contrasts needed to be tested in total in order to account for all degrees of freedom. I coded the four hypotheses of this experiment regarding the learning outcomes as three orthogonal focal contrasts and included four additional, orthogonal residual contrasts. Table 5 lists the three focal contrast codings (representing the four hypotheses regarding learning outcomes) and the four residual contrasts.

In Hypotheses 1 and 3, I postulated that participants in CTRL would show the worst performance (coded -1), whereas participants in MIX would have the best performance (coded +1). The other six experimental groups were expected to show learning performances better than CTRL and worse than MIX (all coded 0). In Hypothesis 2, I expected a main effect of process type in that I expected those participants in groups with integration processes (both groups coded +2) to perform better than those in groups with either only text or picture comprehension processes (all four groups coded -1). CTRL and MIX were coded 0; the difference of these two groups relative to the other groups was already sufficiently described in the first contrast, and since all contrasts are entered simultaneously into the analysis, the relationship between these two groups and the other six groups is already accounted for. It is important to note that all the focal and residual contrasts are orthogonal to each other, meaning that they explain the data while the other contrasts are controlled for. If one of the contrasts explains the data to a significant degree, it does so while simultaneously taking all the other contrasts into account. Finally, in Hypothesis 4, I expected a main effect of the number of implementation intentions, so that the internalization of three implementation intentions (all three groups coded +1) would lead to a better learning outcome than of only one implementation intention (all three groups coded -1). Again, CTRL and MIX were coded 0 for the same reason as explained above.

For the analysis of the gaze data, I also used contrast coding in order to test my hypotheses (Hypotheses 5 through 8). For each of these analyses, I coded one focal contrast according to my hypothesis and six orthogonal residual contrasts. Table 6 lists the focal and residual contrasts for the analysis of the gaze data.

Table 6: Effect coding of focal and residual contrasts for the eye movement analysis in Experiment 2.

	1-TXT	**1-PIC**	**1-INT**	**3-TXT**	**3-PIC**	**3-INT**	**MIX**	**CTRL**
Focal contrast: Fixation time on text	0	-1	0	+1	-1	0	+2	-1
Residual contrast 1	+2	0	-1	0	0	-1	0	0
Residual contrast 2	0	0	+1	0	0	-1	0	0
Residual contrast 3	0	-1	0	0	-1	0	0	+2
Residual contrast 4	0	+1	0	0	-1	0	0	0
Residual contrast 5	0	+1	0	-9	+1	0	+6	+1
Residual contrast 6	-5	+3	-5	+3	+3	-5	+3	+3
Focal contrast: Fixation time on pictures	-1	0	0	-1	+1	0	+2	-1
Residual contrast 1	0	+2	-1	0	0	-1	0	0
Residual contrast 2	0	0	+1	0	0	-1	0	0
Residual contrast 3	-1	0	0	-1	0	0	0	+2
Residual contrast 4	+1	0	0	-1	0	0	0	0
Residual contrast 5	+1	0	0	+1	-9	0	+6	+1
Residual contrast 6	+3	-5	-5	+3	+3	-5	+3	+3
Focal contrast: Transitions between text and pictures	-1	-1	+1	-1	-1	+2	+2	-1
Residual contrast 1	0	0	0	0	0	+1	-1	0
Residual contrast 2	-1	-1	0	-1	-1	0	0	+4
Residual contrast 3	+3	-1	0	-1	-1	0	0	0
Residual contrast 4	0	+2	0	-1	-1	0	0	0
Residual contrast 5	0	0	0	+1	-1	0	0	0
Residual contrast 6	+1	+1	-15	+1	+1	+5	+5	+1

4.2 Results

4.2.1 Control variables

One-factorial ANOVAs were conducted in order to test for differences in prior knowledge between the eight conditions (see Table 7 for means and standard deviations of all control variables as well as learning outcomes). There was no significant difference between conditions with regard to self-reported prior knowledge, $F < 1$, or the multiple-choice test about cell biology, $F(7,152) = 1.74$, $MSE = 477.12$, $p = .10$, $\eta_p^2 = .07$. However, despite randomization, there was a significant difference between groups on the Life Sciences scale of the Test of Basic Scientific Literacy, $F(7,152) = 3.84$, $MSE = 620.64$, $p = .001$, $\eta_p^2 = .15$. Bonferroni post-hoc tests showed that participants in CTRL ($M = 61.67$; $SD = 13.42$) performed significantly worse in the Test of Basic Scientific Literacy than participants in the 1-INT ($M = 77.50$; $SD = 9.21$; $p = .003$) and 3-TXT ($M = 76.88$; $SD = 13.07$; $p = .01$) group. Finally, the eight groups systematically differed in how much time they had spent on the learning phase, $F(7,152) = 2.08$, $MSE = 104.52$, $p = .05$, $\eta_p^2 = .09$. However, Bonferroni post-hoc tests showed no significant differences between groups.

In order to control for these differences between experimental groups, both scientific literacy as well as learning time were included as covariates in the following analyses. Note that learning time was controlled in the following analyses, despite the fact that it might have mediated the effect of the experimental manipulation on learning outcomes, at least partially. Consequently, the inclusion of learning time results in a more conservative testing of the hypotheses.

With regard to learners' achievement motivation, a MANOVA with the four scales of the QCM as dependent variables was conducted to see whether there were differences between the eight experimental groups. There was no significant effect of experimental condition on the different scales of the QCM, $V = .19$, $F(28,608) = 1.01$, $p = .45$. Means and standard deviations for the different scales can be found in Table 7.

Furthermore, in order to test for the homogeneity of regression slopes, I investigated whether learners' achievement motivation interacted with the use of implementation intentions. For the sake of simplicity, I averaged the four scales of the QCM[3]. I then conducted a one-factorial ANOVA with the experimental conditions and the averaged QCM as independent variables and learning outcomes as dependent variable. There was no significant difference between conditions with regard to the effect of implementation intentions, $F(7,144) = 1.45$, $MSE = 0.01$, $p = .19$, $\eta_p^2 = .07$, the averaged QCM, $F(1,144) = 1.79$, $MSE = 0.02$, $p = .18$, $\eta_p^2 = .01$, or the interaction of both, $F(7,144) = 1.31$, $MSE = 0.01$, $p = .25$, $\eta_p^2 = .06$.

Nevertheless, in order to control for a possible moderation of motivational variables on the effect of implementation intentions, all four scales of the QCM were included as covariates in the following analyses concerning the effect of implementation intentions.

4.2.2 Learning outcomes

As noted above in the Methods section (4.1.6), I used effect coding for the analysis of the learning outcomes. I conducted a multiple regression analysis with the learning outcomes as dependent variable and four subsets entered simultaneously as predictors: the centered covariates (scientific literacy and learning time), the centered QCM scales (probability of success, interest, challenge, and anxiety), the three focal contrasts, and four residual contrasts.

The regression model was significant, adj. $R^2 = .15$, $F(13,146) = 3.13$, $p < .001$. The amount of explained variance due to the first subset of predictors (scientific literacy and learning time) was not significant, $\Delta R^2 = .02$, $p = .12$. The explained variance by the second

[3] The scale Anxiety was inverted for this purpose (cf. Experiment 1).

subset (QCM scales) was significant, $\Delta R^2 = .09$, $p = .003$, as was the variance explained by the third subset (focal contrasts), $\Delta R^2 = .07$, $p = .01$. Finally, variance explained by the fourth subset (residual contrasts) was not significant, $\Delta R^2 = .01$, $p = .62$, which means that, in an omnibus test, the focal contrasts unambiguously fitted the data with respect to the hypotheses that they represented, whereas alternative contrasts did not fit the observed data.

Table 7: Means and standard deviations for control variables and learning outcomes in Experiment 2.

Number of implementation intentions	**1**			**3**				
Type of processing	**TXT**	**PIC**	**INT**	**TXT**	**PIC**	**INT**	**MIX**	**CTRL**
Self-reported prior knowledge (1-4)	2.03 (0.57)	2.08 (0.71)	2.15 (0.65)	2.08 (0.94)	2.15 (0.67)	2.05 (0.60)	2.03 (0.75)	1.68 (0.83)
Multiple choice test about cell biology (% correct)	45.77 (15.28)	47.69 (15.30)	46.54 (16.07)	43.85 (12.75)	44.23 (14.31)	42.31 (12.35)	46.15 (15.78)	35.00 (14.65)
Scientific Literacy (% correct)	73.54 (15.90)	64.79 (16.19)	77.50 (9.21)	76.88 (13.07)	70.63 (9.22)	71.25 (9.74)	73.96 (12.74)	61.67 (13.42)
Learning time (in minutes)	20.39 (6.94)	21.64 (7.78)	19.1 (5.31)	19.67 (5.74)	22.56 (10.18)	25.24 (7.97)	24.76 (4.85)	20.54 (6.46)
Probability of success (1-7)	4.51 (1.04)	4.55 (1.09)	4.51 (0.86)	5.11 (1.11)	4.78 (1.25)	4.58 (1.01)	4.44 (1.43)	4.51 (0.82)
Interest (1-7)	3.86 (0.79)	3.89 (1.24)	3.77 (0.98)	4.31 (0.74)	3.89 (0.87)	3.93 (1.02)	4.01 (1.16)	4.16 (0.82)
Challenge (1-7)	4.81 (0.81)	4.79 (1.26)	4.86 (0.99)	4.80 (0.95)	4.96 (0.86)	5.15 (1.01)	4.96 (0.80)	5.40 (0.96)
Anxiety (1-7)	3.50 (1.32)	3.80 (1.56)	3.27 (1.60)	3.13 (1.35)	2.85 (1.36)	4.26 (1.32)	3.67 (1.82)	3.25 (1.45)
Learning outcome (mean accuracy)	.71 (.09)	.72 (.10)	.71 (.08)	.69 (.08)	.67 (.11)	.68 (.11)	.77 (.08)	.66 (.12)

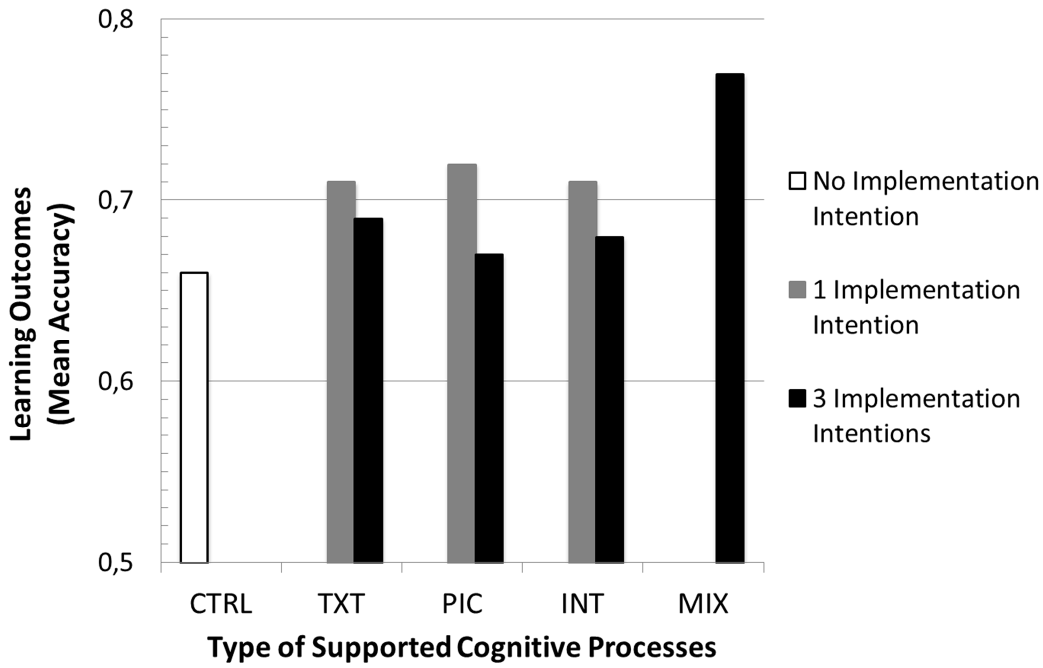


Figure 7. Mean accuracy on the post-test in Experiment 2 (max. 1.0).

Table 8 lists all *B*-, *SE*-, β-, and *p*-values of this analysis. The analysis showed that participants in CTRL performed significantly worse than the other groups, whereas participants in MIX showed significantly better learning outcomes than the remaining groups. With regard to the effect of the type of cognitive process on learning outcomes, there was no significant main effect; that is, there was no difference between the groups whose implementation intentions included text, picture, or integration processes. Finally, there was a significant main effect of the number of implementation intentions. However, as can be seen from the negative β-value, this effect was contrary to my expectation; three implementation intentions about the same process type led to worse performance than a single implementation intention. Figure 7 shows the mean accuracy of the learning outcomes.

Table 8: B-, *SE*-, β-, and *p*-values for learning outcomes in Experiment 2.

		B	***SE***	β	***p***
Learning outcomes	Scientific literacy	0.001	0.001	0.15	.07
	Learning time	0.001	0.001	0.09	.28
	Probability of Success	0.02	0.01	0.24	.02
	Interest	0.01	0.01	0.12	.21
	Challenge	-0.01	0.009	-0.10	.24
	Anxiety	0.01	0.01	0.18	.05
	Focal contrast 1 (CTRL worst, MIX best)	0.04	0.02	0.20	.01
	Focal contrast 2 (TXT, PIC < INT)	-0.001	0.01	-0.02	.82
	Focal contrast 3 (1 < 3)	-0.02	0.01	-0.17	.02
	Residual contrast 1	0.002	0.01	0.02	.84
	Residual contrast 2	-0.01	0.01	-0.04	.64
	Residual contrast 3	0.01	0.004	0.12	.13
	Residual contrast 4	<0.001	0.01	-0.01	.95

4.2.3 Gaze data

Multiple regression analyses with effect coding were conducted with the respective eye-tracking measure as dependent variable and four subsets as predictors: scientific literacy as covariate (centered), the four scales of the QCM (centered), one focal contrast, and six residual contrasts. Since total fixation time constitutes a subset of the overall learning time, I decided to exclude learning time as a covariate in the analysis of gaze data. Table 9 lists the means and standard deviations of the gaze data, while Table 10 lists all *B*-, *SE*-, β-, and *p*-values of the following analyses.

Table 9: Means and standard deviations for gaze data in Experiment 2.

	CTRL	**1-TXT**	**1-PIC**	**1-INT**	**3-TXT**	**3-PIC**	**3-INT**	**MIX**
Total fixation time (seconds)								
Text paragraphs	646.49 (239.71)	607.87 (276.24)	700.20 (339.60)	566.19 (204.30)	636.36 (213.43)	646.02 (282.77)	783.55 (229.59)	702.17 (182.34)
Relevant picture elements	4.44 (1.80)	4.99 (3.10)	7.27 (3.56)	5.51 (2.79)	4.70 (2.10)	7.41 (7.11)	8.26 (3.73)	6.13 (2.64)
Transitions (frequency)								
*Text paragraph to corresponding picture element, vice versa**	2.42 (1.36)	2.99 (1.69)	2.46 (2.09)	3.20 (1.15)	2.83 (1.26)	2.29 (1.20)	3.96 (2.26)	3.95 (1.47)

* For the purpose of data analysis, gaze transitions were added up across all pages of the learning materials.

Experiment 2

Table 10: B-, SE-, β-, and p-values for gaze data in Experiment 2.

		B	***SE***	β	***p***
Total fixation time (text)	Scientific literacy	-670.54	1579.68	-0.04	.67
	Probability of Success	30049.28	24356.75	0.13	.22
	Interest	-48853.96	25777.43	-0.19	.06
	Challenge	71065.08	23103.52	0.27	.003
	Anxiety	6381.04	15449.80	0.04	.68
	Focal contrast	12942.15	20265.33	0.05	.52
	Residual contrast 1	-15040.79	22228.02	-0.05	.50
	Residual contrast 2	-96178.79	39935.17	-0.19	.02
	Residual contrast 3	-16600.29	23043.23	-0.06	.47
	Residual contrast 4	31705.61	39801.97	0.06	.43
	Residual contrast 5	3045.49	5189.60	0.05	.59
	Residual contrast 6	1577.88	5167.99	0.02	.76
Total fixation time (pictures)	Scientific literacy	-13.35	23.17	-0.05	.57
	Probability of Success	604.12	357.28	0.17	.09
	Interest	-444.18	378.12	-0.11	.24
	Challenge	1125.25	338.90	0.28	.001
	Anxiety	96.29	226.63	0.04	.67
	Focal contrast 1	733.05	286.55	0.19	.01
	Residual contrast 1	172.46	338.77	0.04	.61
	Residual contrast 2	-1135.94	585.80	-0.15	.05
	Residual contrast 3	-341.54	351.25	-0.08	.33
	Residual contrast 4	150.85	576.62	0.02	.79
	Residual contrast 5	-118.86	74.67	-0.12	.11
	Residual contrast 6	-185.98	75.13	-0.19	.01

Table 10 (contd.)

Transitions	Scientific literacy	-0.01	.01	-0.10	.221
	Probability of Success	0.07	0.16	0.04	.67
	Interest	-0.25	0.17	-0.14	.14
	Challenge	0.25	0.15	0.14	.09
	Anxiety	-0.11	0.10	-0.10	.26
	Focal contrast 1	.046	0.10	0.36	.0001
	Residual contrast 1	-0.02	0.26	-0.01	.94
	Residual contrast 2	-0.09	0.08	-0.08	.30
	Residual contrast 3	0.11	0.10	0.08	.28
	Residual contrast 4	-0.05	0.15	-0.03	.73
	Residual contrast 5	0.42	0.26	0.12	.11
	Residual contrast 6	0.02	0.02	0.06	.47

For the total fixation time on text, the regression model was only marginally significant, adj. R^2 = .06, F(12,147) = 1.78, p = .06. Of all four subsets, only learners' achievement motivation explained variance to a significant degree, ΔR^2 = .06, p = .04. More specifically, challenge significantly predicted learners' fixation time on the text. Additionally, more interested learners showed a tendency to spend less fixation time on text, as can be concluded by the negative β-value. The other three subsets did not explain variance to a significant degree: scientific literacy, ΔR^2 = .001, p = .67, the focal contrast, ΔR^2 = .002, p = .52, or the residual contrasts, ΔR^2 = .05, p = .27.

Table 11: B-, *SE*-, β-, and *p*-values for the impact of the gaze data on learning outcomes in Experiment 2.

		B	***SE***	β	***p***
Learning outcomes	Scientific literacy	0.001	0.001	0.17	.03
	Probability of Success	0.02	0.01	0.20	.05
	Interest	0.02	0.01	0.16	.08
	Challenge	-0.02	0.01	-0.16	.07
	Anxiety	0.01	0.01	0.20	.02
	Fixation time (text)	<0.001	<0.001	0.04	.64
	Fixation time (pictures)	<0.001	<0.001	-0.02	.83
	Transitions	0.02	0.01	0.26	.002

With regard to the total fixation time on the pictures, the regression model was significant, adj. R^2 = .12, F(12,147) = 2.86, p = .001. The first subset (scientific literacy) did not explain any variance, ΔR^2 = .002, p = .57. However, the other three subsets did explain a significant amount of variance: learners' achievement motivation, ΔR^2 = .07, p = .02, the focal contrast, ΔR^2 = .04, p = .01, the residual contrasts, ΔR^2 = .08, p = .04. Among the QCM scales, challenge significantly predicted fixation time on pictures, while probability of success only had a marginal effect on the fixation time on pictures. Although the hypothesized group differences in the focal contrast could explain the data to a certain extent, so could two of the residual contrasts (residual contrast 2 and 6; see Tables 9 and 10). Residual contrast 2 reflects a higher fixation time on pictures for 1-INT than for 3-INT; this contrast thus matches the unexpected finding that learners with three implementation intentions regarding one type of multimedia processes performed worse in learning outcomes than learners with only one implementation intention. Residual contrast 6 cannot be meaningfully interpreted.

For the analysis of the transitions between text and picture and *vice versa*, the regression model was significant, adj. R^2 = .11, F(12,147) = 2.55, p = .004. Neither was the amount of explained variance due to the first subset (scientific literacy) significant, $\Delta R^2 = .01$, $p = .22$, nor due to the second subset (QCM scales), $\Delta R^2 = .03$, $p = .29$. However, the explained variance by the third subset (focal contrast) was significant, $\Delta R^2 = .12$, $p < .001$. The variance explained by the fourth subset (residual contrasts) was not significant, $\Delta R^2 = .03$, $p = .50$. This means that the focal contrast accurately described the data, meaning that learners in the INT and MIX conditions showed more gaze transitions between corresponding text paragraphs and picture elements and *vice versa*, while 3-INT and MIX showed the most gaze transitions.

Finally, I tested the hypothesis that these three eye-tracking measures are predictive for learning by conducting a multiple regression analysis with learning outcomes as dependent variable and eight predictors that were entered simultaneously: scientific literacy (centered), the four scales of the QCM (centered), and the three eye-tracking measures. Table 11 lists all *B*-, *SE*-, β-, and *p*-values of this analysis. The regression model was significant, adj. $R^2 = .16$, $F(8,151) = 4.67$, $p < .001$. Of the eight predictors, four were significant: As can be seen by the positive β-values, the covariate scientific literacy positively predicted learning outcomes, as did the covariates probability of success and anxiety, as well as gaze transitions. Furthermore, the remaining covariates, interest and challenge, were marginally significant predictors. Interest tended to positively predict learning outcomes, whereas challenge tended to negatively predict learning outcomes. Hence, among the three eye-tracking measures, only the transitions between corresponding text paragraphs and picture elements and *vice versa* significantly predicted learning outcomes in that more transitions yielded better learning.

4.2.4 Post-hoc analyses

The analysis concerning the impact of implementation intentions on learning outcomes had shown that there was a significant effect that groups with three implementation intentions about a single type of multimedia process performed worse than groups with only one implementation intention. A possible explanation of this effect could be that three implementation intentions attracted learners' gazes exclusively towards one representational format, thus causing learners to neglect the other representational format (cf. Wieber & Sassenberg, 2006). To test this assumption, I conducted analyses of covariance (ANCOVAs) for the two TXT conditions and the two PIC conditions (i.e., 1-TXT vs. 3-TXT, 1-PIC vs. 3-PIC). Each ANCOVA either had fixation time on text or fixation time on pictures as dependent variable; scientific literacy (centered) and the four scales of the QCM (centered) served as covariates in these ANCOVAs.

Regarding fixation time on the text for the TXT conditions, there was neither a significant effect of scientific literacy, $F < 1$, nor of number of implementation intentions, $F < 1$. Regarding motivational covariates, there was no significant effect for probability of success, $F < 1$, interest, $F(6,33) = 2.58$, MSE < 0.001, $p = .12$, $\eta_p^2 = .07$, or anxiety, $F < 1$, but there was a significant effect of challenge, $F(6,33) = 8.22$, MSE < 0.001, $p = .01$, $\eta_p^2 = .20$. That is, learners that felt more challenged spent more time on the text (β = 0.48).

For fixation time on pictures for the TXT conditions, there were no significant effects at all, $F < 1$ for scientific literacy, number of implementation intentions, probability of success, interest, challenge, and anxiety, $F(6,33) = 2.21$, MSE < 0.001, $p = .15$, $\eta_p^2 = .06$.

For the PIC conditions, there was no significant effect of scientific literacy or number of implementation intentions on fixation time on text, both $F < 1$. With regard to the motivational covariates, there was no significant effect of probability of success, $F < 1$, interest, $F(6,33) = 1.76$, MSE < 0.001, $p = .19$, $\eta_p^2 = .05$, challenge, $F(6,33) = 1.13$, MSE < 0.001, $p = .30$, $\eta_p^2 = .03$, or anxiety, $F < 1$.

For the fixation time on pictures, there was no difference between PIC conditions with regard to scientific literacy and the number of implementation intentions, both $F < 1$. Among the motivational covariates, only challenge had a significant effect on the fixation time on pictures, $F(6,33) = 6.89$, MSE < 0.001, $p = .01$, $\eta_p^2 = .17$. That is, learners who felt more challenged also spent more time looking at the pictures ($\beta = 0.48$). Probability of success, $F(6,33) = 1.73$, MSE < 0.001, $p = .20$, $\eta_p^2 = .05$, interest, $F < 1$, and anxiety, $F(6,33) = 1.90$, MSE < 0.001, $p = .18$, $\eta_p^2 = .06$, had no significant effect on the fixation time on pictures.

In effect, it does not seem as if the significant inferiority of three implementation intentions regarding one process type was caused purely by an overbearing attentional pull towards one representational format at the expense of the other.

4.3 Discussion

The aim of Experiment 2 was to revisit the question of whether implementation intentions can be used to foster effective processing during multimedia learning. In addition to addressing the general effectiveness of implementation intentions, I tried to answer the question of what type of multimedia processes should best be supported by means of implementation intentions (text processes, picture processes, text-picture integration processes, or a combination thereof) and whether more than one implementation intention could be used concurrently.

As hypothesized, the analyses showed that implementation intentions improved learning outcomes in comparison to the control group (Hypothesis 1). Moreover, there was confirmation of the hypothesis that a widespread support for all three types of effective multimedia processes via three implementation intentions resulted in the best learning outcomes (Hypothesis 3). However, there was neither evidence for the hypothesis that the support of text-picture integration processes would lead to better learning outcomes than the support of either only text or picture processes (Hypothesis 2), nor for the hypothesis that three implementation intentions would always result in better learning than a singular one (Hypothesis 4). Quite to the contrary, three implementation intentions pertaining to one type of multimedia processes actually led to less learning than a single implementation intention.

With regard to the gaze data, the expected impact of implementation intentions on the total fixation time on text was not found (Hypothesis 5). For the total fixation time on pictures, the hypothesized group differences were found, that is, increased fixation times on the pictures for learners in the conditions with implementation intentions pertaining to picture processes, integration processes, and the combination of the three process types (Hypothesis 6). However, next to the focal contrast there were also two significant residual contrasts, only one of which is easily interpreted (residual contrast 2); this residual contrast described a model in which participants in the 1-INT group spent more time looking at the pictures (+1) than participants in the 3-INT group (-1), while the rest of the groups were assumed to not significantly differ from the grand mean (all 0). Residual contrast 2 thus partly matched the finding that learners with three implementation intentions pertaining to one type of multimedia processes had worse learning outcomes than learners with only one implementation intention of the same type. So while the hypothesized pattern of group differences accurately described the data to a certain extent, this finding was not unambiguous because there were additional models that accurately described the data. Only for the transitions between text and pictures, the hypothesis was confirmed in that the groups with implementation intentions about integration processes and the combination of different process types looked back and forth more often between the two representational formats (Hypothesis 7). Interestingly, among the eye-tracking measures, these transitions were the only significant predictor of learning outcomes, which supports findings of Mason and colleagues (2013). Thus, the hypothesis that these three eye-tracking measures would predict learning outcomes was confirmed only partially (Hypothesis 8).

While there were no hypotheses with regard to learners' task-specific learning motivation in Experiment 2, there were significant effects of various scales on learning outcomes and eye-tracking parameters. Since these effects were controlled for in all the analyses, it can be concluded that the effect of implementation intentions is not only dependent on learners' motivational orientation, as was the case in Experiment 1, but occurs also after motivation has been accounted for.

In summary, it can be concluded that implementation intentions can support the use of effective cognitive processes during multimedia learning. In order to maximize the effectiveness of implementation intentions, it is best to use three implementation intentions supporting text, picture, and integration processes at the same time. The analyses of the gaze data suggest that transitions between text and pictures are influenced by the use of implementation intentions and are predictive of learning outcomes.

These findings raise the question why there were no differences in learning outcomes between the three supported types of multimedia processes. Similarly to Experiment 1, the most plausible explanation for this finding is that it is not enough to support only the higher-order integration of text and pictures without supporting the construction of a mental model of the text and the pictures at the same time. According to the theories of multimedia learning, such as the CTML or the ITPC, the successful integration process is preceded by the processing of text and picture information as well as the construction of representationally specific mental models based on this information (Mayer, 2009; Schnotz, 2005). It was implicitly assumed that learners would necessarily process both the text and the pictures in order to successfully deploy the integration implementation intentions. However, this task might have been too ambitious even for implementation intentions; in a series of studies, Dewitte, Verguts, and Lens (2003) found evidence that implementation intentions might work less well for difficult goals if the implementation intentions stress the goal instead of the actions necessary to achieve the goal. Similarly, by stressing the goal of integrating information across text and pictures while skipping the intermediate steps of constructing mental models from the information in the text and pictures first, the implementation intentions about integration processes might actually have been suboptimal for achieving the learning goal.

Why did the groups learning with three implementation intentions perform worse in the post-test than the groups learning with only one implementation intention? One possible explanation for this finding is that three implementation intentions regarding the processing of one representational format might have strongly attracted learners' attention to that representational format, so that they consequently failed to take the other representational format fully into account (for the attention attraction effect of implementation intentions, see Wieber & Sassenberg, 2006). Since the gaze data did not reveal any differences in attention allocation on text and pictures between the conditions with one implementation intention and those with three implementation intentions, however, this explanation seems unsatisfactory. Another possible explanation is that there might have been an interference effect between three comparatively similar implementation intentions. It is unlikely that the implementation intentions were not specific enough to add value to a single implementation intention (cf. De Vet et al., 2011). Instead, learners might have initiated effective multimedia processes but then aborted them as soon as the situational trigger of another, similar implementation intentions was encountered. Moreover, learners might have confused situational triggers and their corresponding actions. In contrast to more complex tasks during which different situational triggers are distributed across larger time intervals, the situational triggers for the implementation intentions in this experiment could be encountered several times per page. The condition with mixed implementation intentions, on the other hand, had only one situational trigger per representational format per page, so that it was possibly easier to distinguish among them and to chain the actions evoked by implementation intentions into more sensible sequences.

Another interesting question is why the groups with implementation intentions regarding text processes, integration processes, and the combination of processes types did not have longer fixation times on the text than the groups with implementation intentions regarding picture comprehension processes? One possible explanation for this finding is that learners already had a differentiated and habitual set of text processes that they used regardless of what type of implementation intentions they received. Furthermore, the text was necessary to understand the learning materials, so that learners had to spend a certain amount of time on reading the text anyway. Finally, learners show a general tendency to focus on the text at the expense of the pictures (Schmidt-Weigand et al., 2010a). These factors together might have resulted in an equal amount of text processing between different groups so that the eye movement measure could not meaningfully differentiate between the experimental conditions.

A point of criticism for Experiment 2 is that, although there are seven experimental conditions, there is only a single control condition that did not receive any implementation intentions or any information about effective multimedia processes at all. Thus, it is not possible to ascertain whether the positive effect of the implementation intention conditions on learning outcomes compared with the control condition was due to the inherent effectiveness of implementation intentions or because the implementation intentions conveyed information about effective multimedia processes to learners. This experimental design was chosen deliberately for practical reasons, despite its limitation. For this reason, Experiment 3 tried to improve on this limitation and investigate the robustness of the results obtained in Experiment 2 at the same time.

5 Experiment 3

Experiment 3 aimed at replicating the main result in Experiment 2 with another, stronger control condition. Control groups in research about implementation intention are usually instructed to internalize a goal intention (i.e., "*I want to X*") instead of an implementation intention. Therefore, I decided to use a strong control condition, in which learners internalized goal intentions about using effective multimedia processes. The control condition contained exactly the same information about effective multimedia processes as the implementation intention condition except that the information were not presented in the format that is characteristic for implementation intentions, namely, the "if-then" format. This is a subtle, albeit important difference since the "if-then" format is thought to be elemental for triggering the underlying mechanisms behind the positive effect of implementation intentions on goal achievement. In particular, the "if" part leads to a heightened activation of the situational cue in memory, which in turn facilitates its recognition, whereas the "then" part establishes the strong link between trigger and action, which in turn enables automatized action. In Experiment 3, I compared a goal intention condition with the best experimental implementation intention condition in Experiment 2, that is, the MIX condition. For this comparison, I expected to replicate the findings from Experiment 2 in that the implementation intention condition should show a better learning performance than the goal intention condition.

5.1 Method

Since Experiment 3 served as a replication for the findings in Experiment 2, albeit with a more conservative control group, its method shared many similarities with the previous experiment.

5.1.1 Participants and Design

42 students from the University of Tübingen participated in this experiment (mean age = 24.21 years, *SD* = 3.25; 29 female). Like previous experiments, biology majors or majors in subjects with a heavy biological focus were excluded from the experiment. Participation was voluntary and was reimbursed with either 15 Euros or course credits. Participants were randomly assigned to one of two experimental conditions with 21 students serving in each condition.

The experiment used a between-subject 2-group experimental design. Participants in the control condition had to internalize the intention to use three effective multimedia processes (goal intention condition), whereas participants in the experimental condition had to internalize three pre-phrased implementation intentions about the use of the same effective multimedia processes (implementation intentions condition).

5.1.2 Materials

5.1.2.1 Learning materials

Participants learned with the same instructional materials as in Experiment 2. For practical reasons, Experiment 3 was not computer-based; instead learners learned with a printed paper-and-pencil version of the same materials.

5.1.3 Measures

Dependent variables for this experiment encompassed participants' learning outcomes; prior knowledge, learning time, and task-specific learning motivation served as control variables.

5.1.3.1 Motivation

Like in previous experiments, participants' motivation was assessed by means of the Questionnaire on Current Motivation (QCM; Vollmeyer & Rheinberg, 2006).

5.1.3.2 Prior knowledge

For assessing learners' prior knowledge, the same three measures like in Experiment 2 were used (see 4.1.3.2): two items assessing learners' self-reported prior knowledge (cf. Mayer & Moreno, 1998), 24 items from the Life Sciences scale of the Test of Basic Scientific Literacy (Laugksch & Spargo, 1996), and a 13-items multiple-choice test on cell biology.

5.1.3.3 Post-test

I used a printed paper-and-pencil version of the same post-test as in Experiment 2 (see 4.1.3.3). In Experiment 3, the post-test had an internal consistency of Cronbach's $\alpha = .63$.

5.1.4 Procedure

The experiment was conducted in small groups with up to eight participants. Participants were tested individually in a single session that lasted about 90 minutes. At the beginning, the experimenter gave all participants a short standardized introduction. Afterwards, participants answered a paper-and-pencil questionnaire for assessing demographic data, their prior knowledge, and current task-specific learning motivation. The questionnaire ended with a concise description of the experiment's structure and domain.

In the goal intention condition, participants received a written instruction to learn with three goal intentions directly before the start of the learning phase. The instruction was phrased as follows (translated from German):

"Please try to learn with the intention to make best use of the learning materials! Therefore, please intend to translate the following intentions into action! *'I will search every picture for its central elements with regard to content!'*, *'I will carefully re-read all paragraphs after having finished a page!'*, *'I will search the picture for the contents described in every paragraph!'*"

Participants in the implementation intention condition received the same implementation intention instruction as the MIX condition in Experiment 2 (see 4.1.5).

After having completed the questionnaire and instruction, participants received the booklet containing the instructional materials. Participants had no time limit during the learning phase. After having finished learning, participants handed in the learning materials, the experimenter noted down participants' learning times, and the participants received the paper-and-pencil post-test and were asked to answer it as accurately and quickly as possible. Finally, they were debriefed and given their remuneration.

5.2 Results

5.2.1 Control variables

Like in Experiment 2, I conducted ANOVAs in order to see whether the two conditions differed with regard to prior knowledge (see Table 12 for means and standard deviations of all control variables and the learning outcome). There were no significant differences between both groups with regard to self-reported prior knowledge, $F(1,40) = 2.17$, $MSE = 2.88$, $p = .15$, $\eta_p^2 = .05$, or the multiple-choice test about cell biology, $F < 1$. Despite randomization, the groups showed a marginally significant difference on the Life Sciences scale of the Test of Basic Scientific Literacy, $F(1,40) = 3.55$, $MSE = 535.71$, $p = .07$, $\eta_p^2 = .08$, in that the group learning with implementation intentions had higher scientific literacy (M = 76.39, SD = 12.24) than the group learning with goal intentions (M = 69.25, SD = 12.32). With regard to time spent on the learning phase, I found no difference between the groups, $F(1,40) = 1.44$, $MSE = 100.60$, $p = .24$, $\eta_p^2 = .04$. In order to rule out that scientific literacy would confound the results of this experiment, I included it as a covariate in the following analysis.

For learners' achievement motivation, a MANOVA with the four scales of the QCM as dependent variables was conducted to see whether there were differences between both experimental groups. There was no significant effect of experimental condition on the different scales of the QCM, $V = .16$, $F(4,36) = 1.77$, $p = .16$. Means and standard deviations for the different scales can be found in Table 12.

As in Experiment 2, I investigated whether learners' achievement motivation interacted with the use of implementation intentions in order to test for the homogeneity of regression slopes. Again, I averaged the four scales of the QCM (see Section 4.2.1). I then conducted a one-factorial ANOVA with the experimental conditions and the averaged QCM as independent variables and learning outcomes as dependent variable. There was no significant difference between conditions with regard to the effect of implementation intentions, $F < 1$, the averaged QCM, $F < 1$, or the interaction of both, $F < 1$.

Nevertheless, in order to control for a possible moderation of motivational variables on the effect of implementation intentions, all four scales of the QCM were included as covariates in the following analyses concerning the effect of implementation intentions.

Table 12: Means and standard deviations for control variables and learning outcome in Experiment 3.

	Goal intention	**Implementation intention**
Self-reported prior knowledge (1-4)	2.62 (1.05)	3.14 (1.25)
Multiple choice test about cell biology (% correct)	38.46 (15.76)	42.49 (12.55)
Scientific Literacy (% correct)	69.25 (12.32)	76.39 (12.24)
Learning time (in minutes)	25.19 (7.21)	22.10 (9.38)
Probability of success (1-7)	4.69 (0.90)	5.06 (0 .91)
Interest (1-7)	3.99 (1.39)	4.09 (1.27)
Challenge (1-7)	4.85 (1.15)	4.58 (0.82)
Anxiety (1-7)	2.82 (1.33)	3.16 (1.57)
Learning outcome (mean accuracy)	.65 (.10)	.73 (.08)

5.2.2 Learning outcomes

A multiple regression analysis was conducted in order to analyze learning performance. The analysis used six predictors, entered simultaneously: the experimental condition (the goal intention condition was coded as -1, the implementation intention condition was coded as +1), scientific literacy (centered), and the four scales of the QCM (centered).

The overall regression model was significant, adj. R^2 = .25, $F(6,34) = 3.20$, $p = .01$. There was a significant effect of experimental condition, in that the group with implementation intentions performed better than the group with goal intentions. Moreover, scientific literacy had a significant effect on learning outcomes. None of the motivational scales were significant predictors for learning outcomes. See Table 13 for all *B*-, *SE*-, β-, and *p*-values of this analysis.

Table 13: B-, *SE*-, β-, and *p*-values for learning outcomes in Experiment 3.

		B	***SE***	β	***p***
Learning outcomes	Experimental condition	0.04	0.02	0.37	.02
	Scientific literacy	0.003	.001	0.39	.02
	Probability of Success	-0.002	0.02	-0.02	.90
	Interest	-0.004	0.02	-0.05	.79
	Challenge	0.02	0.02	0.21	.26
	Anxiety	-0.02	0.01	-0.23	.22

5.3 Discussion

This experiment aimed at replicating the main findings of Experiment 2 against a more conservative control group that received information about effective multimedia processes in form of the explicit goal intention to make use of these processes. Even against this more conservative control group, the implementation intention group achieved significantly higher learning outcomes. This finding not only fits with the robust medium to strong effect of implementation intentions that has been shown in meta-analyses (Gollwitzer & Sheeran, 2006), it also proves the validity of the findings in Experiment 2. Moreover, it suggests that the effect of implementation intentions depends on the specific phrasing as if-then plans; thus, the underlying mechanism is more specific than what could be achieved simply by an instruction to deploy certain cognitive processes (Gollwitzer & Sheeran, 2006).

6 Experiment 4

The previous three experiments showed that implementation intentions can support the use of effective cognitive processes during multimedia learning, thereby improving learning. Experiment 1 indicated that motivation can play an important moderating role on the effect of implementation intentions. Both Experiments 2 and 3 established the positive effect of implementation intentions on multimedia learning regardless of motivational influences and gave a clearer picture of how implementation intentions should be implemented. The purpose of Experiment 4 was to finally contrast implementation intentions with other, more traditional means of supporting the use of effective multimedia processes, such as instructional prompts (e.g., Bannert, 2009; Kombartzky et al., 2010). Although instructional prompts generally follow the *WWW&H* rule and therefore often successfully support self-regulated learning, they have not always been shown to work in multimedia learning. One possible reason might be that executing and keeping up the prompted actions proves still be too cognitively demanding when the instrinsic and extraneous cognitive load of the learning task are already high (e.g., Bartholomé & Bromme, 2009). In fact, under certain conditions, instructional prompts might even add additional extraneous cognitive load by themselves (Horz et al., 2009). At the same time, implementation intentions also generally follow the *WWW&H* rule, while actions elicited by implementation intentions are similar to automated action, meaning that they work even under challenging circumstances, such as high cognitive load (Brandstätter et al., 2001; Gawrilow & Gollwitzer, 2008). Therefore, the following experiment aimed at comparing implementation intentions to explicit instructional prompts and to simple information about the processes with regard to their effect on multimedia learning outcomes under different conditions of cognitive load. Once again, gaze data were assessed as process measures in the following experiment.

Experiment 4

In accordance with research on cognitive load (e.g. Sweller et al, 1998), I expected that all groups learning under conditions of high cognitive load would perform worse than those groups learning under conditions with low cognitive load (Hypothesis 1).

Moreover, I expected that the level of cognitive load would interact with the type of support for using effective multimedia processing that learners received. More specifically, under low levels of cognitive load, I expected a general main effect of instructional support (Hypothesis 2), that is, that explicit instructional prompts or implementation intentions regarding the usage of effective multimedia processes would aid learning compared with a control condition in which learners received only information on these processes (e.g., Kombartzky et al., 2010; Experiment 1 and 3). Under conditions with high cognitive load, however, I expected another pattern in learning outcomes: Even when supported by prompts, learners' self-directed reactions to these prompts still require cognitive resources (Bartholomé & Bromme, 2009; Horz et al., 2009), whereas actions evoked by implementation intentions share similarity with automatized behavior and thus are effective regardless of cognitive load (Brandstätter et al., 2001; Gollwitzer & Sheeran, 2006); consequently, only the internalization of implementation intentions concerning effective multimedia processes should improve learning compared with information on these processes or explicit instructional prompts to use them (Hypothesis 3).

With regard to eye movements, I expected that the three cognitive processes supported in this experiment would be reflected in learners' gaze behavior. Based on prior research (Mason et al., 2013; Experiment 2), I assumed that the transitions between text paragraphs and corresponding picture elements (and *vice versa*) would act as a good indicator for the integration of text and pictures and thus for learning. If the experimental condition has an impact on learning outcomes, this impact should find expression in the number of gaze transitions between text and picture as an indicator for a stronger integration of information from the text and pictures. Accordingly, experimental conditions should have an impact on learners' gaze behavior in the same way that they do on learning outcomes. That is, all groups learning under conditions of low cognitive load should show more gaze transitions than those groups learning under conditions with high cognitive load (Hypothesis 4). Additionally, under a low level of cognitive load, I expected that instructional prompts or implementation intentions about the use of effective multimedia processes would result in more gaze transitions compared with a control condition (Hypothesis 5), whereas under a high level of cognitive load, I expected that the group learning with implementation intentions would show more gaze transitions than the groups without information about effective processes or instructional prompts to use them (Hypothesis 6). Finally, like in Experiment 2, I expected gaze transitions to positively predict learning outcomes (Hypothesis 7).

6.1 Method

6.1.1 Participants and Design

120 students from the University of Tübingen participated in this study (mean age = 23.21 years, *SD* = 3.67; 83 female). Like in previous experiments, students of biology, medicine, or of any field heavily focusing on biology were excluded from participating. For their voluntary participation, students were reimbursed with either 12 Euros or course credits. Participants were randomly assigned to one of six experimental conditions.

The study used a between-subjects 2×3 design with cognitive load and the type of instructional support as factors. Two levels of cognitive load during learning were induced by means of a secondary task. Participants either learned under conditions of low cognitive load (easy secondary task) or under conditions with high cognitive load (difficult secondary task). As the second factor, learners' instructional support was manipulated by giving them different types of instruction prior to learning with multimedia. Participants either received a sheet of paper listing effective multimedia learning processes ("control condition"), an explicit instructional prompt to use the multimedia learning processes ("prompt condition"), or the instruction to use implementation intentions concerning the multimedia learning processes ("implementation intention condition").

6.1.2 Materials

6.1.2.1 Learning materials

The learning materials were the same as in Experiment 2 and 3 (see 4.1.2). The materials were presented on a computer screen with a resolution of 1280×1024 pixels using the E-Prime Professional software by Psychology Software Tools® (version 2.0).

6.1.2.2 Secondary Task

The secondary task was an adaptation of the preload task from Gyselinck, Ehrlich, Cornoldi, de Beni, and Dubois (2000) and Kruley, Sciama, and Glenberg (1994). During the learning phase, learners in all experimental conditions had to remember the position and characteristics of four stimuli displayed in a 4×4 matrix (for an example of the secondary task, see Figure 8). Each stimulus matrix contained two squares, one of them colored, and two sequences of either three digits (e.g., " 718") or three letters (e.g., "MKJ"), one of them colored. These types of stimuli were chosen in order to equally load all slave systems of working memory: the digit/letter sequences were supposed to load the phonological loop, which is responsible for the processing of verbal information, whereas information regarding position and color was assumed to load the visuo-spatial sketchpad, which is responsible for the processing of visuo-spatial information, such as, shapes, colors as well as spatial or movement information (cf. Baddeley, 1992). These stimulus matrices were displayed for eight seconds before they disappeared, during which time learners were instructed to memorize all information contained in them. Later, learners were presented a comparison matrix that either had not been altered compared with the original or had been altered in one of the following ways: 1) one stimulus changed its position within the matrix by one cell, 2) the color of similar stimuli had been changed (i.e., between the two squares or between the two sequences), or 3) the sequence of letters/digits had changed in one stimulus (e.g., "MKJ" to "MJK"). Learners had to rate each comparison matrix either as "correct" or "incorrect".

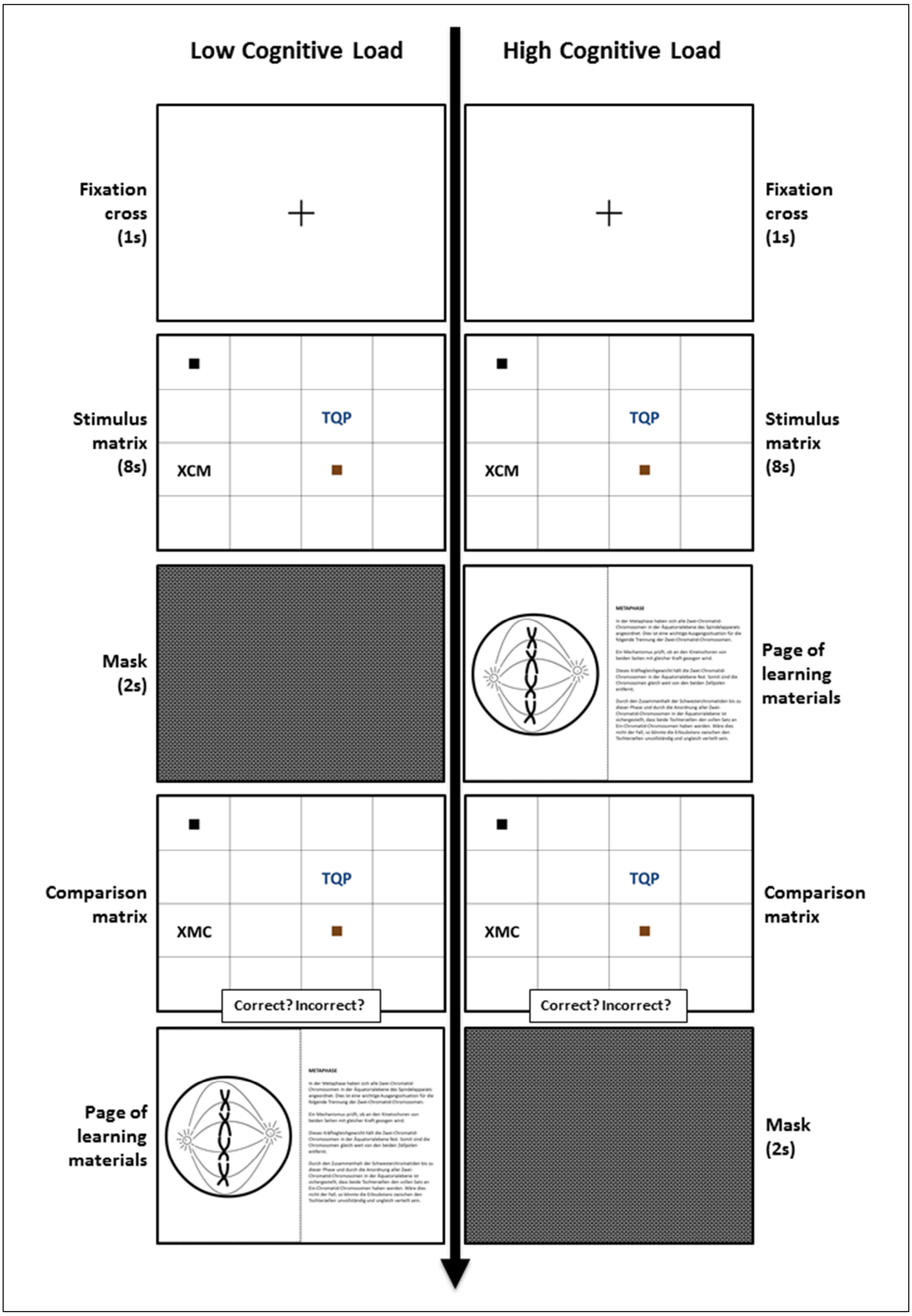
In order to vary the cognitive load during learning, the order of presentation differed between conditions with low and high cognitive load. In conditions with low cognitive load, a stimulus matrix was shown for eight seconds, followed by a mask for two seconds after which the comparison matrix appeared. After each comparison matrix, a page of the learning phase was displayed. Once learners had decided to continue to the next page of the learning materials, the next stimulus matrix was shown. Thus, the secondary task and the primary task (i.e., learning) could be accomplished independently form each other by alternating between the two tasks, thereby causing only little interference. In the conditions with high cognitive load, a stimulus matrix was shown for eight seconds, followed by one page of the learning phase. Once learners had decided to continue to the next page of the learning materials, the comparison matrix was shown. Then, there was a mask for two seconds after which the next stimulus matrix was shown, and so on. Thus, in conditions with high cognitive load, a stimulus matrix had to be maintained in working memory while processing a page of the learning materials. That is, learners were required to work on the primary and secondary tasks in parallel rather than alternating between them, thereby causing stronger interference between the two tasks. Regardless of presentation order, before all stimulus matrices, there appeared a fixation cross for one second in order to draw participants' attention to the center of the screen. All groups had to memorize the same stimulus matrices; only the order of presentation changed between groups with low cognitive load and high cognitive load. The presentation order of matrix pairs (i.e., stimulus and comparison matrices) was randomized for every participant. Like the learning materials, the secondary task was presented via the E-Prime software (version 2.0).

6.1.3 Measures

Dependent variables for this study encompassed participants' learning outcomes, secondary task performance, as well as their self-reported cognitive load. Prior knowledge, learning time, and task-specific learning motivation were assessed as control variables.

6.1.3.1 Motivation

Learners' task-specific learning motivation was assessed by the QCM (Vollmeyer & Rheinberg, 2006; see 3.1.2.2).



,

Figure 8. Presentation sequence of the secondary task for conditions with low and high cognitive load.

6.1.3.1 Prior knowledge

The same three measures as in Experiment 2 and 3 were used to assess learners' prior knowledge (see 4.1.3): Two items for learners' self-reported prior knowledge concerning mitosis and meiosis (cf. Mayer & Moreno, 1998), 24 items from the Life Sciences scale in the Test of Basic Scientific Literacy (Laugsch & Spargo, 1996), and a 13 item multiple-choice test about cell biology.

6.1.3.2 Secondary task performance

For every comparison matrix (see 6.1.2.2), learners had to decide whether it was the same one as the stimulus matrix ("correct") or whether it had changed somehow ("incorrect"). Learners gave their responses via a response box. Each correct response was awarded one point, and the sum of all points was divided by the number of items (mean accuracy).

6.1.3.3 Self-reported cognitive load (SCL)

In order to test whether the manipulation of cognitive load was successful, three 7-point Likert-type scale items were used to measure learners' self-reported cognitive load (SCL) directly after the learning phase ("How much did you concentrate during learning?", "How difficult was the learning content for you?", and "How difficult was it for you to relate the picture with the text?", 1 = "very little" to 7 = "very much"; cf. Cierniak, Scheiter & Gerjets, 2009). For the purpose of data analysis, the three SCL items were averaged.

6.1.3.4 Post-test

Learning outcomes were measured by the same 55 verification items that had been used in Experiment 2 and 3 (see 4.1.3). The post-test was presented via the E-Prime software (version 2.0). It had an internal consistency of Cronbach's $\alpha = .58$.

6.1.3.5 Gaze data

In order to analyze the eye movement data, areas of interest (AOIs) were defined on every page of the learning materials. AOIs included every text paragraph (i.e., every idea unit), and every picture element that was referred to in the text.

All three of the initially investigated eye-tracking parameters (total fixation time on text, total fixation time on picture, gaze transitions between text and picture; cf. Experiment 2) showed a highly significant correlation, $p < .001$. Due to the high multicollinearity between these parameters, only one eye movement parameter was chosen for further analyses: I used the number of gaze transitions between a text paragraph and its corresponding picture element and vice versa. In Experiment 2, the number of gaze transitions had differed between conditions according to my hypotheses and had significantly predicted learning outcomes (corresponding transitions; cf. Johnson & Mayer, 2012; Mason et al., 2013). Hence, when selecting one out of the three predictors due to multicollinearity, learners' gaze transitions were the logical choice.

6.1.4 Apparatus

Learners' eye movements were recorded as process measures during the learning phase using a SensoMotoric Instruments (SMI) High-speed 1250 eye-tracking system with a sampling frequency of 500 Hz. The eye-tracking system was controlled via the SMI iViewX™ software (version 2.4), while the experiment was presented via the E-Prime software (version 2.0). Data analysis was conducted using the SMI BeGaze™ software (version 3.0) and by using a high-speed event detection algorithm with a peak velocity threshold of 40°/s and a minimum fixation duration of 50 ms.

6.1.5 Procedure

Participants were tested individually in a single session that lasted about 120 minutes. At the beginning, participants received a short standardized introduction by the experimenter.

Afterwards, participants' demographic data, their prior knowledge, and current learning motivation were assessed by means of a computer-based questionnaire. The questionnaire concluded with a concise, written description of the experiment's procedure and domain.

Participants were then handed out a paper-and-pencil instruction, depending on the experimental condition (see Table 14 for the exact instruction). They were also shown the first page of the learning materials for a short time so that they had an impression of what to expect from the learning phase.

In the list and in the prompt conditions, participants were asked to attentively read through the instruction twice. In the implementation intention conditions, three pre-phrased implementation intentions were listed at the end of the instruction. Participants were asked to copy these listed, pre-phrased implementation intentions twice by hand.

<u>Experiment 4</u>

Table 14: The instruction for the list, prompt, and implementation intention conditions in Experiment 4.

List
Learning strategies are behavioral techniques that support the processing and encoding of information. The following learning strategies are helpful to make best use of the learning materials! • To search for the central elements in the pictures • To re-read all text paragraphs at the end of a page • To search the picture for the elements described in the text
Prompt
Learning strategies are behavioral techniques that support the processing and encoding of information. The following learning strategies are helpful to make best use of the learning materials! Therefore, we ask you to use the following learning strategies in this experiment: • Please search for the central elements in each picture! • Please re-read all text paragraphs thoroughly at the end of each page! • After each text paragraph, please search the picture for the elements described therein!

Table 14 (contd.)

Implementation intention
Have you ever experienced that you strongly intended to do something but did not do it in the end? This happens often to almost everybody! There is something, however, that can help you in these cases! Implementation intentions are very specific "if-then" plans that are supposed to help you translate an intention (e.g., "I want to go to the gym!") into an action. For having an implementation intention, you link a condition under which you want to perform an action with said action. *Example: "If I come home after work on Friday, then I take my gym bag and visit the gym!"* In the course of this experiment, please learn with the intention and the implementation intention to make optimal use of the learning materials! Implementation intentions can help you to make best use of the learning materials! Therefore, please copy the following implementation intention(s) twice and imagine how you will realize the implementation intention(s) later in the learning phase! It is important that you really want to realize the implementation intention! • If I am looking at a picture, then I will search for its central elements with regard to content! • If I have finished reading a page, then I will carefully re-read all paragraphs! • If I have read a paragraph, then I will search the picture for the contents described therein!

After they had finished reading the instruction (and, if in the implementation intentions condition, copying the implementation intentions twice by hand), they were seated at the eye-tracker. At first, participants received an instruction regarding the secondary task, followed by a test run (one stimulus and comparison matrix pair) in order to familiarize them with the secondary task. Then, the eye-tracking system was calibrated. Following the calibration, the learning phase started. There was no time limit during the learning phase in that learners could progress at their own pace; however, they were not allowed to go back to previous pages. After the learning phase, participants answered the three self-reported cognitive load (SCL) items and the computer-based post-test. They had no time limit for the post-test but were instructed to work as quickly and accurately as possible. Finally, they were debriefed and given their respective remuneration.

6.1.6 Data analyses

For the manipulation check of the secondary task, the learning outcomes, as well as the eye movement data analysis, I analyzed the data via effect coding (see 4.1.6; cf. Abelson & Prentice, 1997; Niedenthal et al., 2002). I had directed hypotheses about how the experimental manipulation would affect the learning outcomes and other dependent variables, such as the secondary task performance, the SCL, as well as their eye movements, and effect coding allowed us to test these directed hypotheses in a straightforward fashion and with minimal loss of statistical power.

Since there were six different groups in this study, five contrasts needed to be tested in total in order to account for all degrees of freedom. I coded the hypotheses of this experiment regarding the learning outcomes as one focal contrast and included four additional, orthogonal residual contrasts. Among the conditions with an easy secondary task, that is, with low cognitive load, I expected the list group to perform worse (coded 0) than both the prompt and the implementation intention groups (both coded +1). Learners in conditions with a difficult secondary task, that is, with high cognitive load, were expected to generally perform worse than learners in the conditions with low cognitive load. I expected the list and prompt groups with high cognitive load to perform worst among all groups (both coded -1), whereas the implementation intention group with high cognitive load (coded 0) should perform at about the same level as the list group with low cognitive load. Stated differently, under low cognitive load, both support measures should be helpful, whereas under high cognitive load, only implementation intentions should be effective. In addition to this focal contrast, I coded four orthogonal residual contrasts. Table 15 lists all contrast codes for all dependent variables used in this experiment.

The design of this study rested on the assumption that I would be able to impose cognitive load in learners by means of the secondary task. I expected the difficult secondary task to impact learners' behavior in two ways: First, learners in conditions with the difficult secondary task should show worse recognition of changes between matrices than learners in conditions with the easy secondary task. Secondly, I expected learners in conditions with the difficult secondary task to self-report higher levels of cognitive load than learners in conditions with the easy secondary task.

In order to test these assumptions, I coded both of these hypotheses as focal contrasts with four additional, orthogonal residual contrasts. Regarding the performance on the secondary task, I expected that the three groups with easy secondary task would show better recognition of changes in matrices (coded +1) than the three groups with the difficult secondary task (coded -1). Concerning the SCL after the learning phase, I expected an inverted pattern, that is, that the three groups with the easy secondary task would show lower SCL (coded -1) whereas the three groups with the difficult secondary task would report higher SCL (coded +1).

Table 15: Effect coding of focal and residual contrasts for secondary task outcomes, self-reported cognitive load, and learning outcomes in Experiment 4.

	Low cognitive load			High cognitive load		
	List	**Prompt**	**Impl. intention**	**List**	**Prompt**	**Impl. intention**
Secondary task						
Focal contrast	+1	+1	+1	-1	-1	-1
Residual contrast 1	0	-1	+1	0	0	0
Residual contrast 2	0	0	0	-1	+1	0
Residual contrast 3	-2	+1	+1	-1	-1	+2
Residual contrast 4	-2	+1	+1	+1	+1	-2
Self-reported cognitive load						
Focal contrast	-1	-1	-1	+1	+1	+1
Residual contrast 1	0	0	0	0	-1	+1
Residual contrast 2	-1	+1	0	0	0	0
Residual contrast 3	-1	-1	+2	-2	+1	+1
Residual contrast 4	+1	+1	-2	-2	+1	+1
Learning outcomes, gaze data						
Focal contrast	0	+1	+1	-1	-1	0
Residual contrast 1	+1	0	0	0	0	-1
Residual contrast 2	0	0	0	-1	+1	0
Residual contrast 3	+2	-3	+1	-1	-1	+2
Residual contrast 4	+2	+2	-4	-1	-1	+2

For the analysis of gaze transitions, I expected the same group differences as with the learning outcomes. Under low cognitive load, I expected the list group to show fewer gaze transitions (coded 0) than the prompt or the implementation intention groups (both coded +1). Under high cognitive load, I expected the list and prompt groups to show the fewest gaze transitions between text and pictures among all groups (both coded -1), whereas learners in the implementation intention group should show an average number of gaze transitions between text and pictures (coded 0).

6.2 Results

6.2.1 Control variables

ANOVAs were conducted in order to test for differences in prior knowledge (see Table 16 for means and standard deviations for all control variables). There were no significant differences between conditions with regard to self-reported prior knowledge, $F < 1$, the multiple-choice test on cell biology, $F(5,114) = 1.24$, $MSE = 317.65$, $p = .30$, $\eta_p^2 = .05$, or on the Life Sciences scale of the Test of Basic Scientific Literacy, $F(5,114) = 1.24$, $MSE = 214.47$, $p = .29$, $\eta_p^2 = .05$. Moreover, the six groups did not systematically differ with regard to learning time, $F < 1$.

Concerning group differences between learners' task-specific learning motivation, I conducted a MANOVA with the four scales of the QCM as dependent variables. There was no significant effect of experimental condition on the different scales of the QCM, $V = .10$, $F < 1$, $p = .92$. Table 16 reports the means and standard deviations for the different scales. Like in the two previous experiments, there was no systematic difference between experimental groups regarding learner's achievement motivation.

As in the previous experiments, I tested whether learners' achievement motivation interacted with the experimental conditions (see Section 4.2.1). I conducted a one-factorial ANOVA with the experimental conditions and the averaged QCM as independent variables and learning outcomes as dependent variable. There was no significant difference between conditions with regard to the effect of implementation intentions, $F < 1$, the averaged QCM, $F < 1$, or the interaction of both, $F < 1$.

Nevertheless, in order to control for a possible moderation of motivational variables on the effect of implementation intentions, all four scales of the QCM were included as covariates in the following analyses concerning the effect of implementation intentions.

6.2.2 Secondary task performance and cognitive load

First, I conducted a multiple regression analysis with secondary task performance as dependent variable and the contrasts (one focal contrast, four residual contrasts, see Table 15 for all contrast codings) as predictors. The predictors were entered in two subsets: the focal contrast as a first subset and the four residual contrasts as a second subset. The regression model was significant, adj. $R^2 = .31$, $F(5,114) = 11.45$, $p < .001$. The amount of explained variance due to the first subset (focal contrast) was significant, $\Delta R^2 = .31$, $p < .001$, whereas the explained variance due to the second subset (residual contrasts) was not significant, $\Delta R^2 = .02$, $p = .41$. Thus, learners in conditions with the difficult secondary task were less able to recognize changes between matrices than learners in conditions with the easy secondary task. The *B*-, *SE*-, β-, and *p*-values of this analysis and the following analyses are listed in Table 17.

Afterwards, I conducted another multiple regression analysis with SCL as dependent variable and with the contrasts (one focal contrast, four residual contrasts) as predictors. Again, the focal contrast was entered as a first subset, followed by the four residual contrasts as another subset. The regression model was not significant, adj. $R^2 = .03$, $F(5,114) = 1.67$, $p = .15$. However, the variance that was explained by the first subset (focal contrast) was significant, $\Delta R^2 = .03$, $p = .04$. The second subset (residual contrasts) did not significantly explain variance, $\Delta R^2 = .03$, $p = .39$. Since the focal contrasts tests very specific group differences, the significant effect of the focal contrast can be interpreted despite the regression model generally not describing the data much better than a standard model. Thus, these results indicate that the learners in conditions with difficult secondary task did self-report a higher degree of cognitive load during learning than learners in conditions with easy secondary task.

Overall, these results suggest that the experimental manipulation of inducing cognitive load in learners by means of a difficult secondary task had been successful, since learners in conditions with the difficult secondary task not only performed objectively worse on the task itself but also rated their own cognitive load higher during learning than learners in conditions with the easy secondary task.

6.2.3 Learning outcomes

A multiple regression analysis was conducted with the learning outcomes as dependent variable and three subsets as predictors: the focal contrast as the fist subset, the four scales of the QCM as the second subset, and the four residual contrasts as the third subset. The regression model was significant, adj. $R^2 = .09$, $F(9,110) = 2.23$, $p = .03$. The focal contrast explained variance in the data to a significant degree, $\Delta R^2 = .05$, $p = .01$, as did the second subset with the four motivational variables, $\Delta R^2 = .09$, $p = .02$. The third subset (residual contrasts) was not significant, $\Delta R^2 = .01$, $p = .95$. Table 17 lists all *B*-, *SE*-, β-, and *p*-values of this analysis.

Thus, the hypothesized pattern of group differences could be found in the data; basically, there was a main effect of cognitive load in that learners in conditions with low cognitive load performed better than learners in conditions with high cognitive load. Moreover, there was an interaction between the level of cognitive load and the instructional support that learners received. Under conditions of low cognitive load, learners with prompts or implementation intentions regarding effective multimedia processes outperformed learners receiving only a list of these processes. Under conditions of high cognitive load, however, the advantage of prompts disappeared. Then, prompts resulted in the same level of performance as a list describing multimedia processes, whereas implementation intentions supported the use of effective cognitive processes to such a degree that learners performed as well as those learners in the list condition with low cognitive load (see Figure 9).

Additionally, motivation again proved to be predictor for learning outcomes. More specifically, both probability of success and anxiety were positive predictors for learning outcomes, whereas interest and challenge were not.

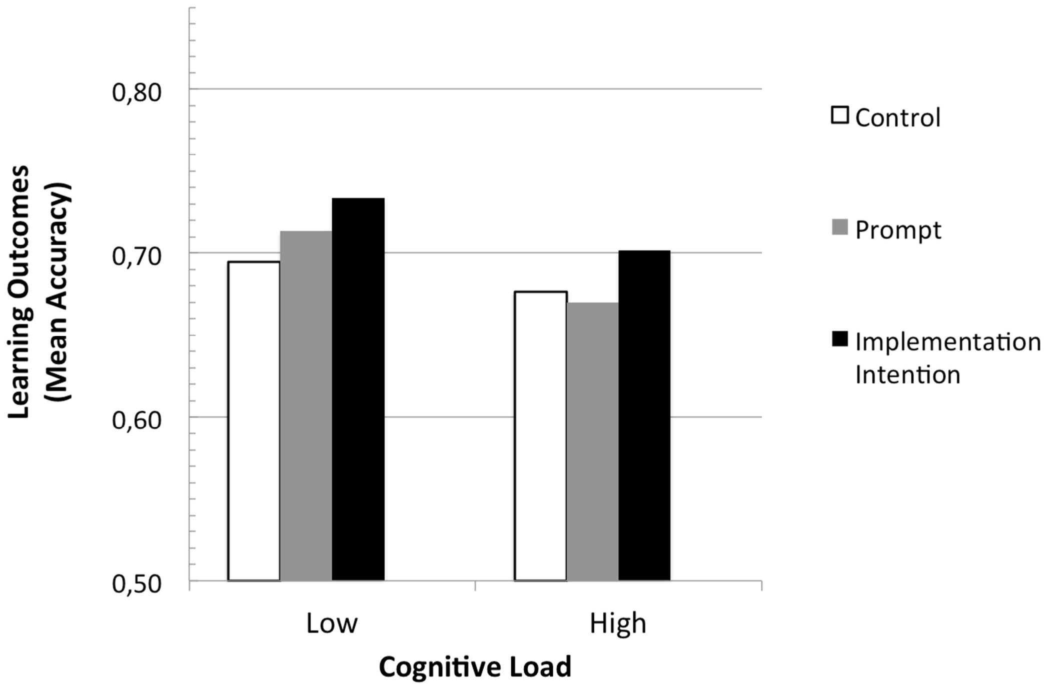


Figure 9: Mean accuracy on the post-test in Experiment 4 (max. 1.0).

Table 16: Means and standard deviations for control variables, cognitive load, learning outcome, and gaze data in Experiment 4.

	Low cognitive load			**High cognitive load**		
	List	**Prompt**	**Impl. intention**	**List**	**Prompt**	**Impl. intention**
Self-reported prior knowledge	1.98	2.08	1.98	2.10	2.20	1.95
(1-4)	(0.73)	(0.71)	(0.60)	(0.85)	(0.68)	(0.84)
Multiple choice test about cell biology	53.08	59.23	49.23	52.31	47.69	51.92
(% correct)	(14.1)	(16.95)	(14.64)	(17.58)	(14.25)	(18.15)
Scientific Literacy	72.71	79.79	75.21	72.08	71.25	76.88
(% correct)	(13.55)	(13.74)	(15.62)	(12.97)	(11.14)	(11.27)
Cognitive Load	5.22	4.90	4.83	5.23	5.48	5.13
(1-7)	(0.87)	(0.63)	(0.77)	(1.00)	(0.83)	(0 .86)
Learning time	23.92	23.95	25.70	21.50	25.84	25.84
(in minutes)	(7.74)	(6.95)	(9.02)	(7.0)	(9.71)	(6.37)

Table 16 (contd.)

Probability of success	4.64	4.73	4.60	4.58	4.45	4.63
(1-7)	(1.43)	(1.36)	(1.40)	(1.23)	(1.14)	(1.23)
Interest	3.94	4.01	4.10	3.41	3.94	4.02
(1-7)	(1.25)	(1.00)	(1.13)	(0.93)	(0.72)	(0.86)
Challenge	4.91	4.88	4.64	4.68	4.89	4.76
(1-7)	(0.90)	(0.97)	(0.62)	(1.12)	(1.05)	(0.86)
Anxiety	3.60	3.38	3.42	3.48	3.76	3.61
(1-7)	(1.38)	(1.19)	(1.79)	(1.53)	(1.21)	(1.57)
Performance on secondary task	.87	.85	.80	.68	.69	.69
(mean accuracy)	(.10)	(.12)	(.10)	(.13)	(.13)	(.12)
Learning outcome	.71	.72	.71	.69	.67	.68
(mean accuracy)	(.09)	(.10)	(.08)	(.08)	(.11)	(.11)
Number of transitions between text and	952.35	953.80	853.20	783.53	999.65	939.70
picture and vice versa	(389.46)	(341.82)	(284.90)	(263.19)	(483.78)	(259.97)

Experiment 4

Table 17: B-, SE-, β-, and p-values for all predictors on secondary task performance, self-reported cognitive load, and learning outcomes in Experiment 4.

		B	*SE*	β	*p*
Secondary task performance	Focal contrast	0.08	0.01	0.59	.0001
	Residual contrast 1	-0.03	0.02	-0.11	.14
	Residual contrast 2	0.004	0.02	0.02	.83
	Residual contrast 3	-0.01	0.01	-0.05	.53
	Residual contrast 4	-0.01	0.01	-0.09	.24
Self-reported cognitive load	Focal contrast	0.16	0.08	0.19	.04
	Residual contrast 1	-0.18	0.13	-0.12	.19
	Residual contrast 2	-0.16	0.13	-0.11	.23
	Residual contrast 3	-0.03	0.05	-0.05	.57
	Residual contrast 4	0.04	0.05	0.08	.41
Learning outcomes	Focal contrast	0.02	0.01	0.23	.01
	Probability of Success	0.02	0.01	0.34	.01
	Interest	0.003	0.01	0.04	.77
	Challenge	-0.01	0.01	-0.09	.37
	Anxiety	0.02	0.01	0.24	.04
	Residual contrast 1	-0.003	0.01	-0.02	.83
	Residual contrast 2	-0.004	0.01	-0.02	.79
	Residual contrast 3	0.002	0.004	0.03	.72
	Residual contrast 4	-0.002	0.003	-0.06	.50
Transitions	Focal contrast	12.71	39.91	0.03	.75
(text-picture and vice versa)	Probability of Success	39.86	35.19	0.15	.26
	Interest	-18.83	42.72	-0.05	.66
	Challenge	-6.76	41.53	-0.02	.87
	Anxiety	47.89	29.07	0.20	.10
	Residual contrast 1	6.07	55.24	0.01	.91
	Residual contrast 2	111.52	57.28	0.19	.05
	Residual contrast 3	-0.59	17.54	-0.003	.97
	Residual contrast 4	15.85	14.30	0.10	.27

6.2.4 Gaze data

By finding the hypothesized differences in learning outcomes, I established that the experimental manipulation did have the expected impact on learning. In order to test whether this effect was mirrored in learners' eye movements, I conducted a multiple regression analysis with gaze transitions between text and corresponding picture elements (and vice versa) as dependent variable and three subsets as predictors: the focal contrast as the first subset, the four scales of the QCM as the second subset, and the four residual contrasts as third subset. The regression model was not significant, adj. $R^2 = .07$, $F < 1$, $p = .51$. Neither did the focal contrast explain any variance in the data, $\Delta R^2 = .001$, $p = .75$, nor the second subset (motivational variables), $\Delta R^2 = .03$, $p = .41$, nor the third subset (residual contrasts), $\Delta R^2 = .04$, $p = .29$. Table 17 lists all *B*-, *SE*-, β-, and *p*-values of this analysis.

Although the omnibus test for all four residual contrasts did not significantly explain any variance, the pattern described in one of the residual contrasts (residual contrast 2) did significantly describe the data. Accordingly, under high cognitive load, learners in the list group showed fewer gaze transitions between text and corresponding picture elements (and vice versa) than the grand mean, whereas learners in the prompt group looked back and forth between text paragraphs and picture elements more often than the grand mean. In effect, under conditions of high cognitive load, learners in the control group showed the fewest number of gaze transitions, whereas learners in the prompt group showed the highest number of gaze transitions.

Finally, in order to test whether the number of transitions actually predicted learning, I conducted a linear regression analysis with learning outcome as dependent variable and gaze transitions as well as the four scales of the QCM as predictors. Table 18 lists all *B*-, *SE*-, β-, and *p*-values of this analysis.

Table 18: B-, SE-, β-, and p-values for gaze transitions and motivational variables on learning outcomes in Experiment 4.

	B	*SE*	β	*p*
Gaze transitions	< 0.001	< 0.001	0.02	.83
Probability of Success	0.02	0.01	0.28	.02
Interest	0.01	0.01	0.09	.45
Challenge	-0.01	0.01	-0.09	.38
Anxiety	0.02	0.01	0.21	.08

Thus, in contrast to findings from prior research (e.g. Johnson & Mayer, 2012; Mason et al., 2013; Experiment 2), gaze transitions did not predict learning outcome at all.

6.3 Discussion

The main goal of Experiment 4 was to compare and contrast implementation intentions with other forms of instructional support, such as the use of explicit instructional prompts or information about effective cognitive processes during multimedia learning. Since it was expected that in most circumstances, instructional prompts would be as effective as implementation intentions, this experiment investigated the difference between implementation intentions and prompts under more extreme circumstances, that is, under conditions of high cognitive load.

The use of instructional prompts in multimedia learning has shown mixed results so far. Some studies could show the positive impact of instructional prompts on learning (Kombartzky et al., 2010; Peeck, 1994; Reinking et al., 1988; Weidenmann, 1988), whereas other studies could not (Bartholomé & Bromme, 2009; Brünken et al., 2005; Drewniak, 1992; Hayes & Readance, 1983). One reason for these mixed results might be that, since the initiation and upkeep of the prompted behavior remains subject to learners' self-regulated capacity, the processing of instructional prompts may actually further strain learners' cognitive resources, thus resulting in cognitive overload (Bartholomé & Bromme, 2009; Horz et al., 2009). This may be especially the case for novice learners or if the intrinsic and extraneous cognitive demands of the learning task are already very high. In contrast, implementation intentions have been shown to work regardless of cognitive load and are particularly helpful when cognitive load is high due to the automaticity of behavior elicited by implementation intentions (Gollwitzer & Sheeran, 2006). Therefore, it was expected that, under conditions of high cognitive load, learners who used implementation intentions about the use of effective multimedia processes would show better learning performance than learners who received prompts to use effective cognitive processes.

The manipulation of learners' cognitive load by means of a secondary task was successful. Not only did leaners in conditions with high cognitive load report a higher subjective, self-reported cognitive load, they also performed worse on the secondary task. Based on these findings, it can be concluded that learners in conditions with the difficult secondary task actually had higher cognitive load during learning and, by extension, that the comparison between implementation intentions and prompts under conditions of low and high cognitive load was actually meaningful.

With regard to learning outcomes, the results mirrored my hypotheses: High cognitive load decreased learning outcomes across the board, confirming Hypothesis 1; moreover, cognitive load interacted with the type of instructional support. Under conditions with low cognitive load, both the prompt and the implementation intention groups outperformed the control group (Hypothesis 2 confirmed), whereas under conditions with high cognitive load, the effectiveness of instructional prompts disappeared, so that the implementation intention group outperformed both the prompt and the control group (Hypothesis 3 confirmed). This finding once again strengthens the results of previous research showing that implementation intentions are especially effective when used under adverse conditions, such as high cognitive load (Branstätter, Lengfelder, & Gollwitzer, 2001), response disinhibition associated with a attention deficit disorder (Gawrilow & Gollwitzer, 2008), or low interest (Experiment 1).

Moreover, it highlights the added value of implementation intentions compared with instructional prompts. In contrast to prompts, implementation intentions are more precise with regard to *when* a good opportunity arises to use effective multimedia processes (cf. *WWW&H* rule; Veenman et al., 2006). Additionally, as has been emphasized before, they are very easy to implement; after all, learners only had to copy them twice by hand in order to internalize them. Therefore, implementation intentions have proven themselves to be advantageous compared with prompts.

One possible criticism of this experiment might concern the way how prompts were implemented. In Experiment 4, the prompt was given to learners only before the learning phase, and learners had no access to the prompt during learning. At the same time, previous research has shown that prompts better improve learning if they are given to learners at the right time during the learning phase and that this improvement is mediated by the use of effective strategic processing (Thillmann et al., 2009). Nevertheless, there are two reasons why implementation intentions would still be superior to prompts, even if the latter were given during rather than prior to learning: First, even if prompts were situated into the learning materials, thereby presenting learners the information about effective cognitive processes at the right time, they would still impose further cognitive load on the learner by leaving the initiation and execution of the prompted actions to learners' self-regulatory capacity (Horz et al., 2009). Second, if they were situated in the learning context, by necessity, they would have to be included in the learning materials, thus making them less economical than implementation intentions, which can be applied independently. Based on these two considerations, implementation intentions represent a favorable substitute for instructional prompts.

The insights gained by learners' gaze data were less conclusive. One of the basic assumptions behind the effect of implementation intentions on multimedia learning is that they support effective cognitive processes, thereby improving learning. In Experiment 2, eye movement data were used as process measures in order to investigate whether the use of implementation intentions actually had an impact on learners' cognitive processes during multimedia learning; although not all measures corresponded with the expected patterns, implementation intentions had the expected impact on gaze transitions between text and pictures. Moreover, gaze transitions were a significant predictor for learning outcomes, a finding that reflected the results of previous studies (e.g., Hegarty & Just, 1993; Johnson & Mayer, 2012; Mason et al., 2013). In Experiment 4, however, learners' gaze transitions did not reflect the experimental manipulation. That is, neither cognitive load nor the type of instructional support did have an effect on the number of gaze transitions between text and pictures (Hypotheses 4 to 6). Moreover, gaze transitions did not act as a significant predictor for learning outcomes (Hypothesis 7). Instead, only one of the residual contrasts (residual contrast 2) matched the data, meaning that under conditions of high cognitive load, learners in the control group made significantly fewer gaze transitions whereas leaners in the prompt group made significantly more gaze transitions than all the other groups. It is difficult to interpret this result but, combined with the findings for the learning outcomes, it can be speculated that a high cognitive load might have caused learners in the control group to overly focus on the text, thus neglecting the integration of information across text and picture, whereas it might have caused learners in the prompt group, striving to follow the directions of the prompt, to fail due to cognitive overload.

As a consequence, this raises the question of why the gaze data had no impact on learning outcomes. One possible explanation is that the range in learning outcomes between groups was comparatively low in Experiment 4 (ranging from .67 to .73 in Experiment 4 vs. from .66 to .77 in Experiment 2). It is possibly that this small statistical variation worked against the regression analysis. Another possible explanation is that in Experiment 4, with its length, complexity, and the secondary task, learners shifted their attention between text and pictures not only in order to integrate the information from both representational formats but also because they had more difficulties with learning in general. Thus, depending on context and possibly individual learner, gaze transitions could represent text-picture integration *or* lack of understanding. The number and variance of gaze transitions was exceptionally high in Experiment 4, especially when compared with Experiment 2. This has partly to do with the fact that I used a more precise eye-tracking system in Experiment 4. Nevertheless, it could also mean that gaze transitions might have captured more than just integration in Experiment 4, thus losing its predictive power for learning outcomes. In the end, the question why gaze transitions did not predict learning outcomes in Experiment 4 remains unanswerable for the time being.

In summary, my hypotheses about the advantage of implementation intentions over other types of instructional support have been supported. Implementation intentions per se represent a good way to foster the use of effective cognitive processes, thereby improving learning. The full potential of implementation intentions becomes obvious under adverse conditions, such as a high cognitive load. Under these circumstances, implementation intentions still continue to support the learner, whereas prompts or information about effective multimedia processes fail in supporting the learner. Especially with regard to the simplicity and ease with which implementation intentions can be implemented, they have proven themselves to be a more effective and economical alternative to instructional prompts.

7 General discussion

After a short recapitulation of this dissertation's theoretical background and aim, I will discuss the results across all four experiments.

The use of multimedia, that is, the combination of verbal and pictorial codes for conveying information, has been shown to positively affect learning outcomes (Mayer, 2003). In the quest for further improving this beneficial "multimedia effect", research has mainly focused on uncovering design criteria that allow our cognitive systems to process multimedia materials in the most efficient way (Anglin et al., 2004; Mayer, 2009). Yet, the design of multimedia materials is only one side of the coin, the other side being learners' capability of using the materials effectively (Pressley et al., 1989). This capability consists of learners' knowledge and use of effective cognitive processes during multimedia learning. Among the myriad of potentially effective cognitive processes, this dissertation focused on a set of nine representative ones derived from a review of the literature: Three processes pertained to text processing, that is, inspecting section headings (e.g., Hyönä et al., 2002), rereading text paragraphs (e.g., Weinstein & Mayer, 1986), and connecting information in one paragraph with information in previous paragraphs (e.g., McNamara et al., 2004); three processes concerned picture processing, that is, studying the picture thoroughly before reading the text (e.g., Eitel et al., 2013), searching for crucial picture components (e.g., Hegarty & Just, 1993), and creating meaningful connections between picture components (e.g., Mayer, 2009); and, finally, three processes pertained to text-picture integration, that is, connecting the information in the text with corresponding information in the picture, vice versa (e.g., Mason et al., 2013), and inspecting the picture after having read the text in order to verify one's own understanding (e.g., Hegarty, 1992). Unfortunately, knowing what cognitive processes to use does not necessarily mean learners will actually use them; in fact, research has shown that learners tend to make too little use of multimedia materials (e.g., Schmidt-Weigand et al., 2010b). In order to compensate for this insufficiency, educational research has investigated the use of instructional prompts. Instructional prompts are explicit instructions that are meant to remind learners of utilizing specific, beneficial learning processes and thereby support learners in initiating them at the right time (Bannert, 2009). Although research on instructional prompts tends to yield positive results (e.g., Thillmann et al., 2009), from the perspective of self-regulated learning, an overreliance on external learning aids might actually impede learners' development of self-regulatory capabilities in the long run (Boekaerts, 1999). Moreover, depending on a variety of factors, such as learners' prior knowledge, instructional prompts can overstrain learners' working memory capacity (Horz et al., 2009). Learners' failure to use effective cognitive processes during multimedia learning can also be conceptualized as a problem of self-regulated learning (i.e., the self-directed application of cognitive strategies, metacognitive knowledge and skills, as well as motivation and volition to achieve a learning goal), and, more specifically, a problem of volition (Corno, 2001). One technique that has consistently proven itself helpful in overcoming volitional obstacles is the use of implementation intentions (Gollwitzer, 1999). Implementation intentions are specific if-then plans that strongly link a situational cue (in the "if" part) to an intended behavioral response (in the "then" part), thus facilitating the automatic initiation of the intended action during an opportune moment. At the same time, the cognitive mechanisms underlying the effect of implementation intentions share similarities to automatization and thus circumvent cognitive load on learners' working memory (Gollwitzer & Sheeran, 2006). Based on these considerations, this dissertation dedicated itself to four overarching research questions that provided a framework for the four experiments in this investigation: First, can implementation intentions support the use of effective cognitive processes in multimedia learning and, in consequence, improve learning? Second, in what way should implementation intentions be used to be most effective for multimedia learning? Third, how do implementation intentions compare with other means to achieve similar goals, more specifically, to instructional prompts? Fourth, what is the relationship between implementation intentions and learners' motivation? In the following, these four research questions will be answered by drawing on the results of all four experiments in this dissertation.

7.1 Can implementation intentions support the use of effective cognitive processes and improve learning?

This research question represented the most central one of this dissertation and was tackled by all four experiments. In Experiment 1, the impact of internalizing two pre-phrased implementation intentions was investigated (studying the picture thoroughly before reading the text and connecting the information in the text with corresponding information in the picture). One group that learned with implementation intentions was compared with a control group that did not learn with them but received the same amount of information about effective multimedia processes. Learners in the experimental group had to write down both implementation intentions five times each. For assessing learning outcomes, two measures were used: recall and transfer performance. With regard to recall performance, there was only a marginally significant effect of implementation intentions overall. Implementation intentions did significantly affect recall performance for learners who were less interested in the learning task (which will discussed in Section 7.4). That is, learners who learned with the two implementation intentions tended to show better recall performance than learners who learned without them. For transfer performance, there was no effect of implementation intentions at all. Since there were no valid process measures in Experiment 1, it is unknown whether learners actually acted upon the two implementation intentions.

In Experiment 2, the main focus laid on the question on how implementation intentions should be implemented effectively (see Section 7.2 for this particular discussion). Overall, nine pre-phrased implementation intentions were used in Experiment 2, either containing effective cognitive processes pertaining to text processing, picture processing, text-picture integration, or a combination of all three. Learners internalized either one or three implementation intentions. The groups learning with implementation intentions were compared with a group that learned without implementation intentions or any information about effective multimedia processes at all. Building on the insights gained from Experiment 1, the instruction for using implementation intentions was slightly changed, so that learners only had to write down the implementation intentions twice instead of five times. For Experiment 2 and the consecutive experiments, a new and more consistent learning measure was developed. Additionally, as a process measure, learners' gaze behavior was recorded via eye-tracking technology. All seven groups that learned with implementation intentions showed better learning outcomes than the control group that learned without implementation intentions. The group learning with three mixed implementation intentions (i.e., one pertaining to text processing, one to picture processing, and one to integration) showed the best learning outcomes amongst all groups. Although learners' gaze behavior was not easily mapped on the cognitive processes evoked by implementation intentions, there were indications that implementation intentions supported certain cognitive processes, such as the integration of corresponding text and picture components. These cognitive processes then improved learning.

Since learners in the control group of Experiment 2 did not receive any information about effective multimedia processes at all, Experiment 3 tried to replicate the results of Experiment 2 with a more conservative control group. The condition with the best learning outcomes in Experiment 2 (the mixed condition) was compared with a control group that learned with goal intentions about using effective multimedia processes. The result of Experiment 2 could be replicated, that is, the group learning with implementation intentions showed better learning outcomes than the group learning with goal intentions. In particular, results from this experiment suggest that the implementation intention effect is more specific than just resulting from telling learners which processes to deploy during learning, which was also part of the instruction in the control condition. Not only that, goal intentions also include a motivational component (“I want to do…”), so the effect of using implementation intentions must have been based on more than just motivation. Thus, implementation intentions trigger specific processes due to their formulation as if-then plans, which may also explain why they have been shown to have rather enduring effects (cf. Gollwitzer & Sheeran, 2006).

In Experiment 4, the experimental condition with the best learning outcomes in Experiment 2 (the mixed condition) was compared with a control condition that received only information about effective multimedia processes. At the same time, a dual-task paradigm was used to induce either low or high cognitive load during learning. Again, learners’ gazes or, more specifically, learners’ gaze transitions between corresponding text and picture components were assessed as a process measure. As hypothesized, there was a main effect of cognitive load on learning outcomes so that learners with low cognitive load learned better than learners with high cognitive load. Moreover, regardless of learners’ cognitive load, the implementation intention groups showed significantly better learning outcomes than the groups that received only information about effective multimedia processes. In contrast to Experiment 2, however, learners’ gaze transitions were not a significant predictor for learning outcomes, which may have to do with the low range in learning outcomes between groups working against the regression analysis or learners having more difficulties with learning so that gaze transitions might have represented either text-picture integration *or* lack of understanding (cf. Section 6.3). Therefore, it is unknown whether implementation intentions actually affected the intended cognitive processes and whether these improved learning subsequently.

General Discussion

To summarize, in three out of four experiments, there was a beneficial effect of learning with implementation intentions. In Experiments 2 to 4, the groups learning with implementation intentions consistently showed better learning outcomes than their respective control groups. The only exception to this pattern was found in Experiment 1, in which the effect of implementation intentions was limited to learners with lower interest and to recall performance, whereas there was no effect of implementation intentions on transfer performance.

As discussed in Section 3.3, there are several possible explanations for the only weak effect of implementation intentions at an overall level: The fact that implementation intentions only had a marginal effect on recall performance might be traced back to the instruction on how to internalize implementation intentions. Having to write down each implementation intention five times each might have resulted in learners' reactance and annoyance (especially for those who were initially motivated to learn). The same explanation also works for learners' transfer performance. In addition, it should be pointed out that the transfer items had a low internal consistency which might explain why there was no effect; simply put, the items might have been too unreliable for a positive effect to show. Alternatively, in addition to the implementation intention instruction, the lack of effect on transfer performance might be explained by a suboptimal choice of implementation intentions and the multimedia processes contained therein. That is, supporting only a picture process and an integration process might not have been enough without supporting text processing also. On the one hand, this second explanation seems little convincing in light of the results in Experiment 2 in which even the groups learning with one implementation intentions pertaining to text-picture integration outperformed the control condition. On the other hand, it should be remembered that, in contrast to Experiment 1, the control condition in Experiment 2 did not receive any information about effective multimedia processes at all. Hence, this alternative explanation should not be discarded outright.

Judging the results as a whole, the research question of whether implementation intentions can improve multimedia learning can be answered positively based on the data gathered in the present dissertation. Whether this improvement truly resulted from implementation intentions affecting learners' cognitive processes during multimedia learning cannot be determined conclusively. Although implementation intentions had the hypothesized impact on learners' gaze transitions in Experiment 2, and learners' gaze transitions were a significant predictor for learning outcomes, there was no equivalent evidence in Experiment 4. As already discussed in Section 6.3, two possible explanations come to mind for this: One reason could be that there was a small range in learning outcomes in Experiment 4, which might have worked against the regression analysis. Another explanation could be that the manipulation of learners' cognitive load might have increased the instances of learners shifting their attention from text to picture (and vice versa) due to a lack of understanding instead of text-picture integration. Consequently, gaze transitions would have ceased to be a valid indicator for integration processes alone, causing the muddled results.

Overall, the findings of this dissertation fall in line with the majority of research about implementation intentions that could show their beneficial effect on action initiation in a wide variety of contexts, from laboratory to field settings, and with a multitude of dependent variables (Gollwitzer & Sheeran, 2006). They also add to the growing number of studies that show a positive effect of implementation intentions in an educational context (e.g., Duckworth et al., 2011; Oettingen et al., 2000). At the same time, this dissertation showed the effectiveness of implementation intentions in yet another context, that is, the application of cognitive strategies during learning. It demonstrated that implementation intentions are successfully applicable for cognitive processes that learners often neglect to use in a targeted and deliberate fashion.

7.2 How should implementation intentions be used to be most effective?

The second research question flows naturally from the first one. If implementation intentions can indeed support multimedia learning, then how should they best be used for maximum effect? That is, which multimedia processes should be evoked by implementation intentions and how many implementation intentions should be used at the same time? My first experiences with using implementation intentions in Experiment 1 showed mixed results, so consequently this research question was mainly approached by Experiment 2: Three groups learned with one implementation intention (one pertaining to a text process, to a picture process, or an integration process), whereas four groups learned with three implementation intentions (three pertaining to text processes, three to picture processes, or three to integration processes; additionally a mixed group with one implementation intention pertaining to text processes, one to picture processes, and one to integration).

Since multimedia materials can only unfold their full potential when learners actually integrate information across text and picture, it was assumed that the groups learning with implementation intentions about text-picture integration would have better learning outcomes than the groups that learned with implementation intentions about text or picture processing. However, this hypothesis did not find support in the data; in fact, there was no difference between implementation intentions about a single type of multimedia process, they were all equally beneficial for learning. Only when learners learned with a combination of all three, that is, in the mixed group, there was a synergy effect for learning. The mixed group showed the best learning outcomes overall. It seems like it is not enough to support only the lower-order processes of constructing a mental model of the text or the picture or to support only the higher-order process of text-picture integration. Only by supporting all involved processes equally, implementation intentions can achieve maximum effect.

The hypothesis that three implementation intentions about the same type of cognitive process would generally yield better learning outcomes than only one implementation intention could not be confirmed either; on the contrary, learning with one implementation intention proved to be more effective than learning with three implementation intentions pertaining to the same type of multimedia process. Yet, as already mentioned above, the mixed group outperformed all other groups despite learning with three implementation intentions. As discussed in Section 4.3, one plausible explanation for this result is that, in the groups with three implementation intentions about one type of cognitive process, the situational triggers in the “if” part might have interfered with each other; that is, actions evoked by one implementation intention might have been prematurely aborted in favor of actions evoked by another implementation intention. In the mixed condition, however, this would have happened less often since the situational triggers were not encountered multiple times while processing one representational format. Additionally, the mixed group would have benefited from supporting the whole range of cognitive processes involved in multimedia learning instead of only one of them.

The results of Experiment 2, as well as the continued success of the mixed condition in Experiments 3 and 4, therefore suggest that neither the type of cognitive process nor the number of implementation intentions alone are main determining factors for how well implementation intentions work. Rather, both proved to be relevant in their own right. The fact that the mixed group showed the best learning performance leads to the conclusion that one should put much thought into what cognitive processes are really necessary for successful learning and how they are interrelated; based on these considerations, the number of implementation intentions should then be determined. While doing so, it seems to be important to pay special attention to the different situational triggers in the “if” part so that, hopefully, there is as little potential for interference between these triggers as possible. So, based on Experiment 2, the recommendation should be to use as many implementation intentions as necessary with as little trigger overlap as possible. These results somewhat reflect the findings of De Vet and colleagues (2011), who could show that the specificity of implementation intentions were a good predictor for their effectiveness but not their number; more implementation intentions were only more effective when sufficiently specific.

In a way, the attempt to answer the question of how implementation intentions should be best used, even in the specific context of multimedia learning, was entering new ground. With some notable exceptions (Achtziger et al., 2008; De Vet et al., 2011), to my knowledge, there has been very little research that investigated these types of questions with regard to implementation intentions.

7.3 How do implementation intentions compare with instructional prompts?

After having answered the questions whether implementation intentions can support multimedia learning and how, the third research question of this dissertation prompted itself: What added value do implementation intentions really have over other, already established instructional methods such as instructional prompts? Although it is always good to have more options available to support multimedia learning, there would be little impetus to use implementation intentions instead of instructional prompts if they did not prove more effective in some way; one such advantage of implementation intentions over instructional prompts is that the use of instructional prompts can result in cognitive overload (e.g., Horz et al., 2009), whereas implementation intentions are more efficient regarding cognitive ressources (e.g., Brandstätter et al., 2001).

General Discussion

The aim of Experiment 4 was to explore the question of how implementation intentions compare with instructional prompts. Learners either learned with instructional prompts about effective multimedia processes that were presented before the learning phase or with implementation intentions (the mixed condition from Experiment 2). At the same time, a dual-task paradigm was used to induce either low or high cognitive load during learning. Research had already established that instructional prompts can strain learners' cognitive resources, especially if learners are novices (e.g., Horz et al., 2009), whereas responses evoked by implementation intentions share similarities with automatized actions and are thus effective regardless of cognitive load (e.g., Brandstätter et al., 2001). Building on these findings, it was hypothesized that there would be an interaction between instructional method and cognitive load: That is, instructional prompts and implementation intentions should be equally beneficial for learning under conditions of low cognitive load. Under conditions of high cognitive load, however, this beneficial effect should disappear for instructional prompts but remain for implementation intentions. The results of Experiment 4 could clearly confirm this hypothesis: Implementation intentions proved to be effective regardless of cognitive load, whereas instructional prompts completely lost their effectiveness when learners suffered from high cognitive load.

Thus, implementation intentions do compare favorably to instructional prompts. Not only does learning with implementation intention avoid the problem of cognitive overload, it is also completely independent on the learning materials itself and could have a positive effect beyond the learning environment that uses instructional prompts. As such, implementation intentions represent a more flexible and less context-sensitive instructional method than the use of instructional prompts.

7.4 What is the relationship between implementation intentions and motivation?

Finally, the fourth and last research question of this dissertation concerned the relationship between the use of implementation intentions and learners' task-specific learning motivation. The question was mainly focused on in Experiment 1 and then only controlled for in the following experiments.

The base assumptions in Experiment 1 was that, if learners have little motivation to perform well in a task, they would have to exert more volitional control to compensate for their lack of motivation. This type of volitional control is difficult and thus takes a lot of effort. Implementation intentions, however, have shown to be especially helpful when volitional control is difficult (e.g., Orbell & Sheeran, 2000; Sheeran & Orbell, 2000); therefore, they were hypothesized to help less motivated learners in particular. In line with this reasoning, although there was no main effect of implementation intentions on recall performance in Experiment 1, there was an interaction between learners' interest in the learning task and the use of implementation intentions, so that learners with less interest in the learning task significantly benefited from the use of implementation intentions, whereas learners with more interest did not. For transfer performance, there was no interaction between implementation intentions and motivation. So, in effect, the hypothesized interaction was found, but only for one out of four motivational factors (i.e., interest, the rest being probability of success, challenge, and anxiety) and only for recall performance.

In general, learning motivation, as assessed by the Questionnaire on Current Motivation (QCM; Vollmeyer & Rheinberg, 2006), did not predict recall performance in Experiment 1. For transfer performance, only probability of success positively predicted learning, whereas interest was a negative predictor paradoxically. In Experiment 1, it became obvious that the relationship between motivation and implementation intentions is likely to turn out complex.

This complexity could not be accounted for in the following experiments without compromising other research questions; as a consequence, it was decided to set aside the question of motivation and instead focus on these other research questions. Nevertheless, motivation was still controlled for in the following experiments. In Experiment 2, learning motivation (i.e., the averaged QCM) did not interact with the use of implementation intentions. Amongst the four motivational variables, probability of success and anxiety were positive predictors for learning outcomes; this finding was somewhat surprising since it had been assumed that anxiety would negatively predict learning. In Experiment 3, again, there was no interaction between implementation intentions and learning motivation (i.e., the averaged QCM). Moreover, none of the motivational variables were significant predictors for learning outcomes. Finally, in Experiment 4, no interaction between implementation intentions and learning motivation (i.e., the averaged QCM) was found. As in Experiment 2, probability of success and anxiety positively predicted learning outcomes.

Taken together, none of these results fall into a consistent pattern. As a matter of fact, the inconsistent impact of learning motivation on learning outcomes calls into question the QCM's validity and reliability as a measurement of motivation. In that case, the results of Experiment 1 might even have been a statistical fluke. Unfortunately, whether the QCM was actually assessing learners' motivation cannot be answered conclusively. Moreover, based on the findings of all four experiments, the relationship between motivation and implementation intentions remains uncertain. As such, it will fall to future research to find more satisfying answers. Not only should there be research with different measurements of motivation, but future research should also address the question by investigating the interaction of implementation intentions with other motivational constructs besides current achievement motivation (e.g., goal setting and orientation, situated interest, intrinsic vs. extrinsic motivation; cf. Schunk, Pintrich, & Meece, 2008).

7.5 Conclusions

This dissertation investigated whether implementation intentions can support the use of cognitive processes in multimedia learning, thereby improving learning. The sum of all results from the four experiments conducted for this dissertation suggest that, yes, implementation intentions are an effective means to foster the use of effective multimedia processes and improve multimedia learning. Moreover, this dissertation could show that implementation intentions in multimedia learning are especially effective if learners use three of them, each pertaining to a different type of multimedia process (i.e., text processing, picture processing, text-picture integration). To generalize, it can be assumed that implementation intentions can be especially effective, if they cover a broad range of multimedia processes and are phrased in such a specific way that their triggers do not interfere with each other. Furthermore, this dissertation demonstrated that implementation intentions compare favorably to instructional prompts because they are effective regardless of learners' cognitive load. Finally, although there were indications that implementation intentions might help to compensate for learners' low interest in a learning task in Experiment 1, these results could not be replicated in later studies and should thus be taken with caution until this effect has been shown to manifest consistently in future research.

On the one hand, there are a number of practical implications of this dissertation: First, implementation intentions have now been established as an effective and valid instructional method for facilitating learners' cognitive processing during multimedia learning. That is, the consistency of this dissertation's results strongly confirms that implementation intentions work in the context of multimedia learning. So, in short, implementation intentions *can be used* for supporting multimedia learning.

More than that, implementation intentions are actually a better alternative to instructional prompts (see Experiment 4). Due to the cognitive mechanisms underlying implementation intentions, their effectiveness remains unimpaired by increased cognitive load; rather, the cognitive load that would have resulted from deliberately choosing and following through with effective cognitive processes is circumvented by delegating action control from the self to the implementation intention (Gollwitzer & Sheeran, 2006). Thus, the strong link between situational cue and intended action acts as an automatized behavioral shortcut that, in addition, largely fulfills the *WWW&H* criteria set by Veenman and colleagues (2006). In contrast, instructional prompts are sensitive to cognitive load in that their effectiveness significantly reduces once learners suffer from high cognitive load (cf. Experiment 4; Horz et al., 2009). This may have to do with the fact that learners have to keep the instructional prompts in mind during learning, imposing more cognitive load in addition to the learning task's intrinsic cognitive load. Even if instructional prompts are embedded in the learning materials and presented during learning, the cognitive load of consciously initiating and keeping up the prompted behavior strains cognitive resources that could otherwise be used for better comprehending the learning contents. So, not only can implementation intentions be used for supporting multimedia learning, they actually *should be used* instead of instructional prompts.

Another practical implication is that, since implementation intentions only require a short instruction that is completely independent from the learning materials themselves, they can easily be used with any already existing learning materials in a complementary fashion. While implementation intentions compare favorably to instructional prompts, they do not invalidate the research on design criteria for multimedia learning materials in any way. Rather, they simply approach the challenge of multimedia learning from another perspective, that is, the learners instead of the learning materials. In fact, the effect of implementation intentions should have an additive relationship with the effect of well-designed multimedia materials. On the other hand, in practice, there are plenty of learning materials that were not well-designed in accordance with multimedia research; implementation intentions could represent an instructional "failsafe" to make the best of such suboptimal learning materials.

Finally, the cost-benefit ratio for implementation intentions is very good. They are so easy to use and implement that they seem like an ideal means to support multimedia learning in ecologically valid settings, such as classrooms. Even in a setting full of distractions, it should be enough to instruct students to internalize a set of pre-phrased implementation intentions by writing them down two or three times. Moreover, if introduced skillfully in a classroom setting, students then could use implementation intentions in a more generalized and widely applicable fashion in their everyday lives, thereby improving their overall self-regulatory competence.

At the same time, there are also some theoretical implications of this dissertation: For the research about implementation intentions and volition, this dissertation has shown that implementation intentions can also positively affect cognitive processes during learning. Thus, the effectiveness of implementation intentions has been proven in yet another context. It also shows that implementation intentions not only work on a behavioral level, such as the recovery of activation after an operation (e.g., Orbell & Sheeran, 2000), or on an affective-motivational level, such as shielding the course of action from negative inner states (e.g., Achtziger et al., 2008); this dissertation adds to the findings that implementation intentions also work in activating complex cognitive processes that are more ephemeral and difficult to access. This is especially interesting since the findings indicate that a set of well-considered, complementary implementation intentions can set in motion an intricate chain of cognitive processes that positively affects learners' understanding.

For the field of multimedia research, this dissertation tried to reconceptualize multimedia learning as a specific form of self-regulated learning by showing that effective cognitive processes could be activated by means of a volitional technique, resulting in improved learning. In accordance with Boekaerts' (1999) model of self-regulated learning, implementation intentions not only affect the regulation of processing modes (i.e., the choice of cognitive processes) but also the regulation of the learning process themselves. In contrast to instructional prompts, which support learners' metacognition by reminding them of effective cognitive processes, implementation intentions also affect learners' monitoring concerning good opportunities to use effective cognitive processes as well as the initiation and execution of these processes. By doing so, implementation intentions allow learners to spend more cognitive resources on the monitoring, evaluation, and correction of their understanding. Moreover, the knowledge and use of implementation intentions themselves represent a metacognitive skill that could be generalized across a multitude of contexts.

That multimedia learning can be understood in general terms of self-regulated learning is hardly new (e.g., cognitive-affective theory of learning with media, CATLM; Moreno, 2007; Moreno & Mayer, 2007); however, despite a few attempts to connect these fields, the research on multimedia and on self-regulated learning have remained largely unconnected so far. Only recently, the focus of research has shifted onto the intersection of self-regulation, motivation, and multimedia learning (cf. Park, Plass, & Brünken, 2014). In the same vein, this dissertation represents an attempt at bridging this theoretical and empirical divide as well as a reminder of the necessity do so more strongly in the future. The same goes for linking the research on motivation and volition to cognitive research in education in general. Although research on motivation in education continues to be an active field (for an overview, see Schunk, Pintrich, & Meece, 2008), it mostly remains within its own conceptual confines. Moreover, research on motivation is splintered into many separate subfields (e.g., goals and goal-setting, interest, intrinsic and extrinsic motivation, etc.). To my knowledge, there have been relatively few attempts to investigate the impact of motivation on cognitive processes or on the choice thereof.

7.6 Limitations and future research

Considering all four experiments, there were two major limitations of this dissertation. The first and most conspicuous one was that the role of motivation in relation to implementation intentions still remains unexplained. A big part of the problem lies with the Questionnaire on Current Motivation (Vollmeyer & Rheinberg, 2006) which was used to measure learning motivation. On the one hand, the inconsistent relationship between learning motivation, as measured by the QCM, and learning outcomes raises serious questions about the QCM's validity; on the other hand, in all four experiments, participants answered the QCM only once before the learning phase, which might have been not often enough to get an accurate picture of learners' motivation in the course of learning. That is, in Experiment 1, a single time of testing might not have been enough to account for the ebb and flow of motivation during a longer learning task, leading to inconsistent results. As a consequence of these results, which suggested a far more complex relationship between motivation and implementation intentions than initially expected, the following three experiments did not put any focus on this issue anymore; instead learning motivation was effectively reduced to another control variable. That is, by deliberately shifting the focus away from the role of motivation on implementation intentions, the QCM might actually not have been given a "fair chance" to prove its validity in the consecutive experiments.

Yet, this question of how motivation interacts with implementation intentions remains an important issue. For instance, Koestner and colleagues (2002) found that implementation intentions are more effective when used to achieve self-concordant, that is, intrinsically motivated goals than with extrinsically motivated goals. Thus, motivational aspects like goal setting or intrinsic vis-à-vis extrinsic motivation seem to be important factors for how well implementation intentions work; this is especially important if implementation intentions are used as a learning aid because learners often are not intrinsically motivated to learn. However, to my knowledge, there has been little research, both theoretically as well as empirically, on how implementation intentions exactly interact with different conceptions of motivation (e.g., self-determination theory, expectancy-value theory). Therefore, future research should address this issue more thoroughly.

The second limitation of the experiments in this dissertation was that, while the impact of implementation intentions on learning outcomes was hypothesized to be mediated by the cognitive processes contained therein, there is no certainty about this mediation based on the results of Experiment 2 and 4. In fact, the hypothesized mediation built on an untested assumption: On basis of the eye-mind hypothesis (Just & Carpenter, 1980), gaze data were used as process measures for learners' allocation of attention and, in turn, of their cognitive processes. However, gaze data are always difficult to interpret without proper triangulation with other (process) measures. The untested assumption was that the multimedia processes chosen in this study would result in the hypothesized gaze patterns. Yet, they did not for the most part; in Experiment 2, only gaze transitions were affected by implementation intentions in the hypothesized way. At the same time, the fact that gaze transitions positively predicted learning outcomes in Experiment 2 but not Experiment 4 shows that the meaning of the gaze measure is very sensitive to context. Therefore, future research should first establish a clear link between cognitive processes and the process measures, possibly by using several complementary process measurements concurrently. That is, only by unambiguously demonstrating that implementation intentions about effective multimedia processes affect learners' cognitive processing and cognitive processing then improves learning outcomes, an accurate mediation analysis will become possible.

Besides these two limitations, this dissertation raises a lot of questions for future research. For instance, for experimental reasons, the experiments in this dissertation all had a similar procedure and used the same learning materials. Based on the broad applicability of implementation intentions (Gollwitzer & Sheeran, 2006), there is no reason to assume that the effect of implementation intentions in multimedia learning cannot be generalized to other learning materials, but in order to be certain, future research should try to replicate the findings in this dissertation with a variety of multimedia learning tasks. In the same vein, the effect of implementation intentions on multimedia learning should be tested in an ecologically more valid setting, such as actual classrooms, to see how implementation intentions fare outside controlled laboratory conditions. In order to ascertain the superiority of implementation intentions over instructional prompts, Experiment 4 should be replicated with situated prompts (cf. Thillmann et al., 2009).

The question of what implementation intentions should best be used for multimedia learning also represents a promising field of research. The implementation intentions in this dissertation were pre-phrased due to the preselection of effective cognitive processes. However, research about implementation intentions often allows participants to phrase their own implementation intentions. If implementation intentions are used in a less specific learning context, then it might be helpful to let learners design their own implementation intentions according to their own goals in order to increase the effectiveness of the implementation intentions. Another interesting angle for the choice of implementation intentions is whether the implementation intention are supposed to activate discrete behavior in learners or defend the learning process from distractions; although this dissertation could show that the activation of effective multimedia processes by means of implementation intentions works, Gollwitzer and Schaal (1998) found that implementation intentions are actually more effective for protecting a course of action from a tempting distraction. Thus, the effectiveness of implementation intentions for shielding multimedia learning from distractions should be investigated and compared with implementation intentions for activating multimedia processes. Finally, future research could try to identify effective combinations of implementation intentions, either with regard to effective cognitive processes or even crossing the boundary between cognitive, motivational-affective, and/or behavioral processes. Just as the combination of implementation intentions pertaining to text processing, picture processing, and integration was the most successful, it seems plausible that a combination of implementation intentions pertaining to cognition, metacognition, motivation and affect, as well as outward behavior would be the most effective. All in all, implementation intentions in the context of (multimedia) learning offer a wide and promising field for future research.

8 Summary

This dissertation set out to answer the question whether implementation intentions can support cognitive processing during multimedia learning and improve learning. The use of multimedia (i.e., the combination of text and pictures) has been shown to positively affect learning outcomes (Mayer, 2003). This so called 'multimedia effect' depends, among other things, on learners' knowledge and use of effective cognitive processes during multimedia learning. Yet, knowledge of effective cognitive processes does not necessarily lead learners to make use of them; accordingly, learners often fail to effectively use multimedia materials, resulting in suboptimal learning (e.g., Schmidt-Weigand et al., 2010b). The failure to use effective cognitive processes can be conceptualized as a problem of volition (Corno, 2001). One technique that has consistently proven itself helpful in overcoming volitional obstacles is the use of implementation intentions (Gollwitzer, 1999). Implementation intentions are specific if-then plans that strongly link a situational cue to an intended action, thus facilitating automatic action initiation during good opportunities to act.

Four experiments were conducted to investigate the effectiveness of implementation intentions in multimedia learning. All experiments shared a similar procedure: First, learners received either received either instructions to learn with implementation intentions about certain multimedia processes or some other instruction, followed by a learning phase about the topic of cell division. After the learning phase, learners had to answer a learning test. In Experiments 2 and 4, eye-tracking was used as a process measure of learners' attention distribution.

Experiment 1 investigated whether implementation intentions can foster effective processing during multimedia learning, thereby improving learning. Furthermore, it tried to shed light on the question of how learners' task-specific motivation interacts with the use of implementation intentions. The use of implementation intentions had only marginal impact on

learners' recall and no impact on their transfer performance. However, for recall performance, there was an interaction between the use of implementation intentions and learners' interest: Those learners with little interest in the learning task significantly benefited from using implementation intentions, whereas those learners with more interest did not.

Experiment 2 focused again on the question whether implementation intentions can improve multimedia learning by supporting the underlying cognitive processes. Additionally, I was interested in what type of multimedia learning processes or combination thereof should best be supported by the use of implementation intentions. All groups learning with implementation intentions outperformed the control group that learned without implementation intentions. There was no difference in learning outcomes between implementation intentions evoking different types of multimedia process (i.e., text processing, picture processing, or integration). Contrary to my hypothesis, three implementation intentions about a singular process type resulted in worse learning outcomes than just one implementation intention, possibly due to interference between similar implementation intentions. The group that learned with three implementation intentions about different types of processes (i.e., one pertaining to text processing, one to picture processing, and one to integration) had the best learning outcomes.

Experiment 3 aimed at replicating the main finding of Experiment 2 against a more conservative control group learning with goal intentions. The implementation intention group outperformed the goal intention group. Thus, the findings of Experiment 2 could be replicated.

Finally, in Experiment 4, I tried to further delineate the differences between the use of implementation intentions and other effective ways to support the use of multimedia processes, more specifically the use of instructional prompts (e.g., Kombartzky et al., 2010; Thillmann et al., 2009). In order to do so, I studied the effect of both the use of

implementation intentions and of prompts under different conditions of cognitive load. Under conditions of low cognitive load, both the instructional prompt group and the implementation intentions group showed better learning outcomes than the control group. Under conditions of high cognitive load, however, instructional prompts lost their effectiveness so that learners' performance did not differ from the control group. Meanwhile, implementation intentions remained effective even under conditions of high cognitive load, so that the implementation intention group demonstrated better learning outcomes than both the control and the instructional prompt group.

The sum of all results from the four experiments conducted for this dissertation suggest that implementation intentions are an effective means to foster the use of effective multimedia processes and improve multimedia learning. It can be assumed that implementation intentions are especially effective if they cover a broad range of multimedia processes and are phrased in such a specific way that they do not interfere with each other. Furthermore, this dissertation demonstrated that implementation intentions compare favorably to instructional prompts because they are effective regardless of learners' cognitive load.

9 Zusammenfassung

Die Zielsetzung dieser Dissertation war die Beantwortung der Frage, ob Vorsätze (*implementation intentions*) die kognitive Verarbeitung beim Multimedialernen unterstützen und dadurch den Lernerfolg verbessern können. Vorherige Forschung hat gezeigt, dass die Nutzung von Multimedia (d.h. einer Kombination von Text und Bildern) den Lernerfolg verbessern kann (Mayer, 2003). Dieser so genannte „Multimedia-Effekt" hängt unter anderem vom Wissen der Lernenden um effektive kognitive Prozesse während des Multimedialernens sowie deren Nutzung ab. Das Wissen um effektive kognitive Prozesse führt jedoch nicht notwendigerweise dazu, dass Lernende diese wirklich nutzen. Dementsprechend wird multimediales Lernmaterial oft nicht effektiv genutzt, was zu suboptimalem Lernerfolg führt (z.B. Schmidt-Weigand et al. 2010b). Die unzureichende Nutzung effektiver kognitiver Prozesse kann dabei als volitionales Problem konzeptualisiert werden (Corno, 2001). Eine Technik, die sich beim Bewältigen volitionaler Hürden konsistent als hilfreich erwiesen hat, ist die Nutzung so genannter Vorsätze (Gollwitzer, 1999). Vorsätze sind spezifische Wenn-Dann-Pläne, die einen situativen Hinweisreiz stark mit einer intendierten Handlung verknüpfen und dadurch die automatische Handlungsinitiierung in günstigen Momenten erleichtern.

Es wurden vier Experiment durchgeführt, um die Wirksamkeit von Vorsätzen beim Multimedialernen zu untersuchen. Allen Experimenten war ein ähnlicher Ablauf gemein: Zuerst wurden die Probanden entweder instruiert, mit Vorsätzen bezüglich bestimmter Multimediaprozesse zu lernen, oder sie erhielten eine andere Instruktion. Darauf folgte eine Lernphase zum Thema „Zellteilung". Nach der Lernphase mussten die Lernenden einen Lerntest absolvieren. In den Experimenten 2 und 4 wurden zusätzlich die Blickbewegungen der Lernenden als Prozessmaß für deren Aufmerksamkeitsverteilung aufgezeichnet.

Zusammenfassung

Experiment 1 untersuchte, ob Vorsätze eine effektive Verarbeitung beim Multimedialernen unterstützen und dadurch das Lernen verbessern können. Des Weiteren sollte der Frage nachgegangen werden, wie die aufgabenspezifische Lernmotivation der Probanden mit der Vorsatznutzung interagiert. Die Nutzung von Vorsätzen hatte nur einen marginalen Effekt auf die Wiedererkennungsleistung der Lernenden und keinen Einfluss auf deren Transferleistung. Allerdings zeigte sich bei der Wiedererkennungsleistung eine Interaktion zwischen der Vorsatznutzung und dem Lerninteresse: Lernende mit wenig Interesse an der Lernaufgabe zogen signifikanten Nutzen aus den Vorsätzen, während Lernende mit höherem Interesse nicht davon profitierten.

Experiment 2 konzentrierte sich ebenso auf die Frage, ob Vorsätze das Multimedialernen dadurch verbessern können, dass sie die zugrundeliegenden kognitiven Prozesse unterstützen. Weiterhin lag das Interesse darauf, welche Multimedia-Lernprozesse oder welche Kombination davon am besten durch Vorsätze unterstützt werden sollten. Alle Gruppen, die mit Vorsätzen lernten, erzielten bessere Lernergebnisse als die Kontrollgruppe, die ohne Vorsätze lernte. Es gab keinen Unterschied bezüglich des Lernergebnisses zwischen den verschiedenen Vorsätzen, die unterschiedliche Arten von Multimediaprozessen anregten (d.h. Textverarbeitung, Bildverarbeitung oder Integration). Entgegen meiner Hypothese wirkten sich drei Vorsätze bezüglich einer einzelnen Prozessart im Vergleich einem einzelnen Vorsatz negativ auf das Lernergebnis aus, möglicherweise aufgrund von Interferenzen zwischen den sehr ähnlichen Vorsätzen. Die Gruppe mit drei Vorsätzen zu unterschiedlichen Prozesstypen (d.h. ein Vorsatz zur Textverarbeitung, einer zur Bildverarbeitung und einer zur Integration) erzielte die besten Lernergebnisse.

Experiment 3 zielte darauf ab, das Hauptergebnis von Experiment 2 mit einer strengeren Kontrollgruppe zu replizieren, die mit Zielabsichten (*goal intentions*) lernte. Die Gruppe, die mit Vorsätzen gelernt hatte, schnitt im Lernerfolg besser ab als die Gruppe, die

mit Zielabsichten gelernt hatte. Somit konnten die Ergebnisse aus Experiment 2 repliziert werden.

Schließlich wurde in Experiment 4 versucht, den Unterschied zwischen der Vorsatznutzung und anderen effektiven Fördermaßnahmen für die Verarbeitung von Multimedia herauszuarbeiten, insbesondere die Abgrenzung zur Nutzung von Prompts (z.B. Kombartzky et al., 2010; Thillmann et al., 2009). Hierfür wurde der Auswirkung von sowohl Vorsätzen als auch Prompts unter unterschiedlicher kognitiver Belastung untersucht. Bei niedriger kognitiver Belastung zeigten die Vorsatz-Gruppe und die Prompt-Gruppe beide bessere Lernergebnisse als die Kontrollgruppe; bei hoher kognitiver Belastung verloren Prompts jedoch ihre Wirksamkeit, so dass sich der Lernerfolg der Prompt-Gruppe nicht mehr von der Kontrollgruppe unterschied. Gleichzeitig blieben Vorsätze auch unter hoher kognitiver Belastung weiterhin effektiv, so dass die Vorsatz-Gruppe bessere Lernergebnisse als die Prompt- und die Kontrollgruppe erzielte.

Die Gesamtheit der Ergebnisse aller vier Experimente, die für diese Dissertation durchgeführt wurden, legt nahe, dass Vorsätze eine effektive Methode darstellen, die Nutzung von effektiven Multimediaprozessen zu unterstützen und dadurch das Lernen zu verbessern. Es kann davon ausgegangen werden, dass Vorsätze besonders dann effektiv sind, wenn sie eine große Bandbreite von Multimediaprozessen unterstützen und so spezifisch formuliert sind, dass sie nicht untereinander interferieren. Überdies hat diese Dissertation gezeigt, dass Vorsätze eine bessere Alternative zu Prompts darstellen, da sie unabhängig der kognitiven Belastung des Lernenden wirksam sind.

10 References

Aarts, H., Dijksterhuis, A. P., & Midden, C. (1999). To plan or not to plan? Goal achievement or interrupting the performance of mundane behaviors. *European Journal of Social Psychology, 29*, 971-979.

Abelson, R. P., & Prentice, D. A. (1997). Contrast tests of interaction hypothesis. *Psychological Methods, 2*, 315–328.

Achtziger, A., & Gollwitzer, P. M. (2008). Motivation and volition during the course of action. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (pp. 272-295). London: Cambridge University Press.

Achtziger, A., Gollwitzer, P. M., & Sheeran, P. (2008). Implementation intentions and shielding goal striving from unwanted thoughts and feelings. *Personality & Social Psychology Bulletin, 34*, 381–93.

Achtziger, A., Bayer, U. C., & Gollwitzer, P. M. (2012). Committing to implementation intentions: Attention and memory effects for selected situational cues. *Motivation and Emotion, 36*, 287-300.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2), 131-152.

Ajzen, I., Czasch, C., & Flood, M. G. (2009). From intentions to behavior: Implementation intention, commitment, and conscientiousness. *Journal of Applied Social Psychology, 39*, 1356–1372.

Anglin, G. J., Vaez, H., & Cunningham, K. L. (2004). Visual representations and learning: The role of static and animated graphics. In D. Jonassen (Ed.), *Handbook of research* *on educational communications and technology* (Vol. 2, pp. 865–916). Mahwah, NJ: Erlbaum.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Combining fading with prompting fosters learning. *Journal of Educational Psychology*, *95*, 774-83.

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development, 56*, 45-72.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Bannert, M. (2003). Effekte metakognitiver Lernhilfen auf den Wissenserwerb in vernetzten Lernumgebungen. *Zeitschrift für Pädagogische Psychologie, 17*, 13-25.

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 4*, 359-375.

Bannert, M. (2009). Promoting self-regulated learning through prompts: A discussion. *Zeitschrift für Pädagogische Psychologie, 23*, 139-145.

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues. In R. Wyer & T. Srull (eds.), *Handbook of Social Cognition.* New York, NY: Lawrence Erlbaum.

Bartels, J. M., & Magun-Jackson, S. (2009). Approach–avoidance motivation and metacognitive self-regulation: The role of need for achievement and fear of failure. *Learning and Individual Differences, 19*, 459-463.

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: what kind of support for whom? *Journal of Educational Psychology, 101*, 282–293.

References

Bayer, U. C., & Gollwitzer, P. M. (2007). Boosting scholastic test scores by willpower: The role of implementation intentions. *Self and Identity, 6*, 1–19.

Berger, J.-L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning and Instruction, 21*, 416-428.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction, 7*, 161-186.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research, 31*, 445-457.

Brandstätter, V., Lengfelder, A., & Gollwitzer, P. M. (2001). Implementation intentions and efficient action initiation. *Journal of Personality and Social Psychology, 81*, 946–960.

Brophy, J. (1999). Toward a model of the value aspects of motivation in education: Developing appreciation for particular learning domains and activities. *Educational Psychologist, 34*, 75-85.

Brünken, R., Seufert, T., & Zander, S. (2005). Förderung der Kohärenzbildung beim Lernen mit multiplen Repräsentationen. *Zeitschrift für Pädagogische Psychologie, 19*, 61-75.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-281.

Cassidy, S. (2004). Learning styles: An overview of theories, models, and measures. *Educational Psychology, 24*, 419-444.

Chapman, J., Armitage, C. J., & Norman, P. (2009). Comparing implementation intention interventions in relation to young adults' intake of fruit and vegetables. *Psychology & Health, 24*, 317-332.

References

Chi, M. T., DeLeeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*, 439-477.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, *25*, 315-324.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*, 149-210.

Corno, L. (1986). The metacognitive control components of self-regulated learning. *Contemporary Educational Psychology, 11*, 333-346.

Corno, L. (2001). Volitional aspects of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 191–226). Mahwah, NJ: Erlbaum.

Corno, L., & Kanfer, R. (1993). The role of volition in learning and performance. *Review of Research in Education, 19*, 301-341.

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *The Journal of the Learning Sciences, 12*, 91-142.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: guidelines for research and design. *Educational Psychology Review, 21*, 113–140.

De Vet, E., Oenema, A., & Brug, J. (2011). More or better: Do the number and specificity of implementation intentions matter in increasing physical activity? *Psychology of Sport and Exercise, 12*, 471–477.

Dewitte, S., Verguts, T., & Lens, W. (2003). Implementation intentions do not enhance all types of goals: The moderating role of goal difficulty. *Current Psychology, 22*, 73–89.

References

Drewniak, U. (1992). *Lernen mit Bildern in Texten. Untersuchung zur Optimierung des Lernerfolgs bei Benutzung computerpräsentierter Texte und Bilder.* Münster: Waxmann.

Duckworth, A. L., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. M. (2011). Self-regulation strategies improve self-discipline in adolescents: Benefits of mental contrasting and implementation intentions. *Educational Psychology, 31*, 17-26.

Eitel, A. (2013). *How processing a picture facilitates learning from text: evidence for scaffolding* (Doctoral dissertation).

Eitel, A., Scheiter, K., & Schüler, A. (2012). The time course of information extraction from instructional diagrams. *Perceptual and Motor Skills, 115*, 677-701.

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.

Entwistle, N., Hanley, M., & Hounsell, D. (1979). Identifying distinctive approaches to studying. *Higher Education, 8*, 365-380.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283-299.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 117–133). New York, NY: Cambridge University Press.

References

Flores, F., Tovar, M. E., & Gallegos, L. (2003). Representation of the cell and its processes in high school students: An integrated view. *International Journal of Science Education, 25*, 269-286.

Gawrilow, C., & Gollwitzer, P. M. (2008). Implementation Intentions Facilitate Response Inhibition in Children with ADHD. *Cognitive Therapy and Research, 32*, 261-280.

Gawrilow, C., Morgenroth, K., Schultz, R., Oettingen, G., & Gollwitzer, P. M. (2013). Mental contrasting with implementation intentions enhances self-regulation of goal pursuit in schoolchildren at risk for ADHD. *Motivation and Emotion, 37*, 134-145.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*, 493–503.

Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology, 73*, 186–199.

Gollwitzer, P. M., & Schaal, B. (1998). Metacognition in action: The importance of implementation intentions. *Personality and Social Psychology Review, 2*, 124–136.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 38, pp. 69–119). San Diego, CA: Elsevier Academic Press.

Gyselinck, V., & Tardieu, H. (1999). The role of illustrations in text comprehension: What, when, for whom, and why? In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 195–218). Mahwah, NJ: Erlbaum.

Gyselinck, V., Ehrlich, M. F., Cornoldi, C., De Beni, R., & Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer Assisted Learning, 16*, 166-175.

References

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.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*, 99-136.

Hayes, D. A., & Readence, J. E. (1983). Transfer of learning from illustration-dependent text. *The Journal of Educational Research, 76*, 245-248.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion, 11*, 101-120.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 1084–1102.

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

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, *22*, 411-430.

Höffler, T. N. (2010). Spatial ability: Its influence on Learning with visualizations — A meta-analytic review. *Educational Psychology Review, 22*, 245-269.

Horz, H., Winter, C., & Fries, S. (2009). Differential benefits of situated instructional prompts. *Computers in Human Behavior, 25*, 818–828.

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: evidence from eye fixation patterns. *Journal of Educational Psychology, 94(1)*, 44–55.

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied, 18*, 178–91.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review, 95*, 163-182.

Koestner, R., Lekes, N., Powers, T. A., & Chicoine, E. (2002). Attaining personal goals: self-concordance plus implementation intentions equals success. *Journal of Personality and Social Psychology, 83*, 231–244.

Kombartzky, U., Plötzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a strategy for learning from animations. *Learning and Instruction, 20*, 424–433.

Kruley, P., Sciama, S. C., & Glenberg, A. M. (1994). On-line processing of textual illustrations in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition*, *22*, 261-272.

Laugksch, R. C., & Spargo, P. E. (1996). Construction of a paper-and-pencil test of basic scientific literacy based on selected literacy goals recommended by the American Association for the Advancement of Science. *Public Understanding of Science, 5*, 331–359.

Lewis, J., Leach, J., & Wood-Robinson, C. (2000). Chromosomes: The missing link − Young people's understanding of mitosis, meiosis, and fertilisation. *Journal of Biological Education, 34*, 189-199.

References

Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*, 837-858.

Lorch, R., Lemarié, J., & Grant, R. (2011). Signaling hierarchical and sequential organization in expository text. *Scientific Studies of Reading, 15*, 267-284.

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 & Education, 60*, 95–109.

Mayer, R. E. (2003). *Learning and Instruction*. Upper Saddle River, NJ: Merrill.

Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

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

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and instruction, 20*, 167-171.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90(2)*, 312–320.

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers, 36(2)*, 222–233.

Meijer, J., & Riemersma, F. (2002). Teaching and testing mathematical problem solving by offering optional assistance. *Instructional Science, 30*, 187-220.

References

Moreno, R. (2007). Optimising learning from animations by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods. *Applied cognitive psychology, 21*, 765-781.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*, 309-326.

Niedenthal, P. M., Brauer, M., Robin, L., & Innes-Ker, Å. H. (2002). Adult attachment and the perception of facial expression of emotion. *Journal of Personality and Social Psychology, 82*, 419-433.

Oettingen, G., Hönig, G., & Gollwitzer, P. M. (2000). Effective self-regulation of goal attainment. *International Journal of Educational Research, 33*, 705–732.

Orbell, S., & Sheeran, P. (2000). Motivational and volitional processes in action initiation: A field study of the role of implementation intentions. *Journal of Applied Social Psychology, 30*, 780–797.

O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1985). The effects of repeated readings and attentional cues on reading fluency and comprehension. *Journal of Literacy Research, 17*, 129–142.

Park, B., Plass, J. L. & Brünken, R. (2014). Cognitive and Affective Processes in Multimedia Learning. *Learning and Instruction, 29*, 125-127.

Parks–Stamm, E. J., Gollwitzer, P. M., & Oettingen, G. (2007). Action control by implementation intentions: effective cue detection and efficient response initiation. *Social Cognition, 25(2)*, 248–266.

Peeck, J. (1994). Enhancing graphic-effects in instructional texts: Influencing learning activities. *Advances in Psychology, 108*, 291-301.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*, 459-470.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.

Pokey, P., & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology, 82*, 41-50.

Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. 5, pp. 89–129). Greenwich, CT: JAI Press.

Pressley, M., Borkwski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research, 13*, 857-867.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J. M. Findlay, R.W. Kentridge, & R. Walker (Eds.),*Eye movement research: Mechanisms, processes and applications* (pp. 3–22). Amsterdam: Elsevier.

Reinking, D., Hayes, D. A., & McEneaney, J. E. (1988). Good and poor readers' use of explicitly cued graphic aids. *Journal of Literacy Research, 20*, 229-247.

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2005). Investigating the presentation and format of instructional prompts in an electrical circuit analysis computer-based learning environment. *IEEE Transactions on Education, 48*, 531-539.

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica, 47*, 57-66.

References

Rouet, J. F. (2006). *The skills of document use: From text comprehension to Web-based learning.* Mahwah, NJ: Erlbaum.

Sanchez, R. P., Lorch, E. P., & Lorch, R. F. (2001). Effects of headings on text processing strategies. *Contemporary Educational Psychology, 26*, 418–428.

Scheiter, K., & Eitel, A. (2010). The effects of signals on learning from text and diagrams: how looking at diagrams earlier and more frequently improves understanding. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 264–270). Heidelberg: Springer.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction, 19*, 481-494.

Scheiter, K., & Van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Applied Cognitive Psychology, 23*, 1209–1214.

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: theoretical and empirical considerations. *Metacognition and Learning, 6*, 83-90.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist, 26*, 299-323.

Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples. *Zeitschrift für Pädagogische Psychologie, 23*, 129-138.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010a). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20*, 100–110.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010b). Explaining the modality and contiguity effects: New insights from investigating students' viewing behaviour. *Applied Cognitive Psychology, 24*, 226–237.

Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*, 101-120.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–69). New York, NY: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning & Instruction, 13*, 141-156.

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). *Motivation in education: Theory, research, and applications*. Upper Saddle River, NJ: Person/Merrill Prentice Hall.

Schunk, D. H., & Zimmerman, B. J. (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Lawrence Erlbaum.

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior, 29*, 1590-1601.

Seufert, T., & Brünken, R. (2006). Cognitive load and the format of instructional aids for coherence formation. *Applied Cognitive Psychology, 20*, 321–331.

Sheeran, P. (2002). Intention—behavior relations: a conceptual and empirical review. *European Review of Social Psychology, 12*, 1–36.

Sheeran, P., & Orbell, S. (2000). Using implementation intentions to increase attendance for cervical cancer screening. *Health Psychology, 19*, 283-289.

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.

Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251-296.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*, 66–73.

Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of "what" to prompt or also "when" to prompt? *Zeitschrift für Pädagogische Psychologie, 23*, 105–115.

Van den Broek, P. (2012). Individual and developmental differences in reading comprehension: Assessing cognitive processes and outcomes. In J. P. Sabatini, E. R. Albro & T. O'Reilly (Eds.), *Measuring up: Advances in how to assess reading ability* (pp. 39-58). Lanham, MD: Rowman & Littlefield.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Van Genuchten, E., Scheiter, K., & Schüler, A. (2012). Examining learning from text and pictures for different task types: Does the multimedia effect differ for conceptual, causal, and procedural tasks? *Computers in human behavior, 28*, 2209-2218.

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*, 95–99.

Van Merriënboer, J. J. G., & Sluijsmans, D. M. A. (2008). Toward a synthesis of Cognitive Load Theory, Four-Component Instructional Design, and self-directed learning. *Educational Psychology Review, 21*, 55–66.

References

Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition and Learning, 6*, 205-211.

Veenman, M. V. J., Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning, 1*, 3–14.

Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction, 10*, 293-309.

Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review, 18*, 239–253.

Walczyk, J. J., Marsiglia, C. S., Bryan, K. S., & Naquin, P. J. (2001). Overcoming inefficient reading skills. *Journal of Educational Psychology, 93*, 750-757.

Walczyk, J. J., Marsiglia, C. S., Johns, A. K., & Bryan, K. S. (2004). Children's compensations for poorly automated reading skills. *Discourse Processes, 37*, 47–66.

Webb, T. L., & Sheeran, P. (2008). Mechanisms of implementation intention effects: The role of goal intentions, self-efficacy, and accessibility of plan components. *The British Journal of Social Psychology, 47*, 373–95.

Weidenmann, B. (1988). *Psychische Prozesse beim Verstehen von Bildern.* Bern & Stuttgart: Huber.

Weidenmann, B. (2001). Lernen mit Medien. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 415-466). Weinheim: PVU.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., Vol. 3, pp. 315–327). New York, NY: Macmillan.

References

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71-88). London: Routledge & Kegan Paul.

Wieber, F., & Sassenberg, K. (2006). I can't take my eyes off of it: Attention attraction effects of implementation intentions. *Social Cognition, 24*, 723-752.

Wieber, F., Von Suchodoletz, A., Heikamp, T., Trommsdorff, G., & Gollwitzer, P. M. (2011). If-then planning helps school-aged children to ignore attractive distractions. *Social Psychology, 42*, 39–47.

Wieber, F., Sezer, L. A., & Gollwitzer, P. M. (2014). Asking "why" helps action control by goals but not plans. *Motivation and Emotion, 38*, 65-78.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8*, 327-353.

Wirth, J. (2009). Promoting self-regulated learning through prompts (Editorial). *Zeitschrift für Pädagogische Psychologie, 23*, 91-94.

Wolters, C. A. (2003). Regulation of Motivation: Evaluating an Underemphasized Aspect of Self-Regulated Learning. *Educational Psychologist, 38*, 189-205.

Wolters, C. A., Yu, S., & Pintrich, P. (1996). The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences, 8*, 211-238.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.

Zwaan, R., A., & Singer, M. (2003). Text comprehension. In A. C. Graesser, M. A. Gernsbacher & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 83–122). Mahwah, NJ: Lawrence Erlbaum.

11 List of figures

12 List of tables