Optimizing Learning with Dynamic and Static Visualizations to Foster Understanding in the Natural Sciences

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. Tim Kühl

aus Neunkirchen/Saar

Tübingen

2011

Tag der mündlichen Qualifikation:	26.07.2011
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Katharina Scheiter
2. Berichterstatter:	Prof. Dr. Stephan Schwan

An dieser Stelle möchte ich mich bei all jenen bedanken, die mich bei der Erstellung dieser Arbeit unterstützt haben, sei es aufgrund inhaltlicher Anregungen, oder aufgrund persönlicher Unterstützung und Aufmunterung.

Im Besonderen sollen hier erwähnt werden:

- Prof. Dr. Katharina Scheiter
- Prof. Dr. Peter Gerjets
- Prof. Dr. Stephan Schwan
- die AG Wissenserwerb mit Hypermedia
- die NG Wissenserwerb mit Multimedia
- die Mitarbeiter des „Fisch-Projekts“
- die Kollegen des IWM und der TU Dresden
- meine Familie
- und natürlich Nina

Recht herzlichen Dank!

Contents

List of Figures

List of Tables

1 Introduction

As the Programme for International Student Assessment (PISA) showed, the performance of German pupils in the Natural Sciences is considered to be in need of improvement. The educational standards in the Natural Sciences require students, among other things, to understand complex interrelations that change over time, and to integrate methods and knowledge from various subject matters like biology, chemistry, or physics with each other. In particular, topics in these domains are characterized by the fact that they require learners to understand how a change in one variable (e.g., availability of water) affects another variable (e.g., rate of plant growth). Unfortunately, students often fail to understand the complex interplay of such changes in Natural Sciences phenomena. Moreover, the content in these areas often appears decontextualized and abstract, which in turn holds the risk of inhibiting students to engage more thoroughly in processing information related to such phenomena (e.g., Hake, 1998; Taasoobshirazi & Carr, 2008; Whitelegg & Edwards, 2002). These problems are especially accentuated in physics, and, as a prototypical example, in Newton's laws of motion, which consist of three physical laws. These laws describe how forces that act upon an object determine its motion. More precisely, the first law states that an object stays in rest, or moves uniformly respectively, as long as no force acts upon the object. The second law states that the size of a force is determined by the product of mass and acceleration (i.e., change in velocity). The third law states that when an object A exerts a force upon another object B (Actio), object B exerts a force upon object A (Reactio) that has the same size, but works in the opposite direction. These laws of motion can, on the one hand, be regarded as abstract, as well as decontextualized, and they rather contradict the observations one makes in real life (e.g., Hake, 1998). Moreover, learners often have difficulties understanding the interrelationship between these laws – for instance, the impact of movement characteristics of an object, and its impact on the changing directions and sizes of resulting forces (e.g., Waltner, Rachel, & Wiesner, 2006; Waltner, Wiesner, & Rachel, 2007).

To investigate on how to reduce problems with which learners are confronted in the Natural Sciences (i.e., understanding of the interplay of changes between variables as well as decontextualized content), Newton's law of motions were chosen in this thesis, since they prototypically comprise these problems.

In an attempt to overcome the problem of decontextualization, it has been recommended to bring the content into a context with reference to the real world (context-based physics, for a review see Taasoobshirazi & Carr, 2008; see also situated learning, Resnick, 1987). In the current thesis, this was done by connecting the abstract physical principles of Newton's law of motion with a real-world phenomenon, namely the undulatory (i.e., wavelike) motion of swimming fish for generating propulsion. Connecting Newton's law of motion with the movement of swimming fish has recently become a topic of interest in physic didactics, and has already been successfully implemented in school lessons (Waltner, Rachel, et al., 2006; Waltner, Wiesner, et al., 2007). Similarly, illustrating these physical principles with reference to fish locomotion has in the meantime also been added to German schoolbooks such as "Duden – Naturwissenschaft und Technik" by Franik and Klose (2007).

However, even after bringing physics in context, the cognitive challenges of understanding physical principles remain. One proposed solution and promising way to support students' understanding for such phenomena might lie in the use of multimedia (i.e., text and visualizations), at which visualizations are used to augment verbal explanations. The topic of learning with multimedia will be introduced in Chapter 2.1. Since it cannot be taken for granted that multimedia instruction will always foster comprehension, Study 1 of the current thesis addressed if adding visualizations to text aids understanding of physical principles underlying undulatory motion compared to learning only from text.

With the increasing availability of computers in education, it has become common to not only show static visualizations, but also dynamic visualizations (e.g., videos or animations). However, even if at first glance learning with such dynamic visualizations might seem to be superior to learning with static visualizations in general, the current status of research shows that such a view might be too simplistic (e.g., Hegarty, 2004; Höffler & Leutner, 2007; Schnotz & Lowe, 2008; Tversky, Bauer-Morrison, & Bétrancourt, 2002). Rather, it needs to be specified why and for which purposes one would expect dynamic visualizations to be beneficial. At this, it is also crucial to uncover features that might diminish the potential of dynamic visualizations. These issues will be explicated in more detail in Chapter 2.2. For instance, although dynamic visualizations might possess high potentials to convey dynamic features like motion, trajectory, or acceleration, they also might impose high processing demands onto learners due to their high degree of visual complexity, since elements within the visualization move at different locations at the same time (e.g., Lowe, 1999, 2003, 2004).

Therefore, when investigating the potential of dynamic visualizations for learning, specific design characteristics of the learning environment need to be taken into account, because dynamic visualizations are not necessarily helpful in themselves. Rather, they may need to be implemented in learning environments, as well as designed in ways that support the processing of the to-be-conveyed information. Particularly if such conditions are met, the potential of dynamic visualizations might unfold, and as a consequence, their instructional effectiveness might be enhanced. However, there is still a lack of research on how to implement learning with dynamic visualizations in ways that it helps in processing the respective information (e.g., Bétrancourt, 2005; Schnotz & Lowe, 2008; Tversky, Heiser, Lozano, MacKenzie, & Morrison, 2008).

Thus, a major goal of this thesis is to investigate different ways of optimizing multimedia instruction in general, and dynamic visualizations in particular compared to static visualizations. Contrary to static visualizations, dynamic visualizations are supposed to possess a high degree of visual complexity. Following from this, there are two problems that might reduce their instructional effectiveness compared to the static visualizations (cf. Schnotz & Lowe, 2008). First, problems resulting from the need to split attention between visualization and text (inter-representational split-attention effect) may be more severe when learning from dynamic visualizations than from static visualizations: Due to heir higher degree of visual complexity, the processing of dynamic visualizations might be hampered, if learners need to switch their attention between processing written text and visualizations. With spoken text on the other hand, learners can focus their attention on the visualizations, and process verbal and pictorial information in parallel (e.g., Ginns, 2005; Schnotz, 2005). Thereby, problems associated with a high degree of visual complexity might become less severe. Hence, one may assume that learners receiving dynamic visualizations might benefit from spoken text to a greater extent than learners receiving static visualizations. The theoretical rationale for these assumptions will be described in more detail in Chapter 4.1, and was examined in Study 2 of this thesis.

Second, even though the handling with the visual complexity of dynamic visualizations might be disburdened by using spoken text, the problem of the visual complexity within dynamic visualizations still remains. To cope with the visual complexity within dynamic visualizations, it has been recently suggested to improve dynamic visualizations by means of cueing, since cueing may guide learners' attention to the most relevant information within the visualization (e.g., Boucheix & Lowe, 2010; de Koning, Tabbers, Rikers, & Paas, 2010a; Kriz & Hegarty, 2007). Therefore, it was investigated in Study 3, if cueing would be beneficial for multimedia instruction, and, because specifically dynamic visualizations are supposed to suffer from a high degree of visual complexity, would be even more beneficial when learning with dynamic as opposed to static visualizations. The rationale for these assumptions will be explained in Chapter 4.2.

To conclude, the main focus of this thesis is to detect design characteristics that are supposed to optimize learning with multimedia instruction, and particularly to optimize the effectiveness of learning with dynamic as opposed to static visualizations. Thereby, first it will be tested if visualizations in general are helpful for the domain at hand (Study 1). Thereafter, it will be investigated how design characteristics that are supposed to counteract problems arising from an inter-representational split-attention effect, as well as from the visual complexity of dynamic visualizations themselves, influence learning with dynamic and static visualizations, respectively.

The three studies that are designed to examine these research questions will be described in more detail in Chapters 3, 5, and 6. A general discussion of these studies will be given in Chapter 7.

2 Learning with Multimedia

When considering how to improve learning, particularly in domains that are hard to imagine and where understanding the interplay of different variables is important (like the domain in the current study), one often proposed and advocated solution is the use of multimedia instruction. Multimedia can be defined – in a simple, but widely accepted form, with the aim of covering a wide range of research – as the combined presentation of text and visualizations, independent of the medium in which the information will be conveyed (Mayer, 2001, 2005b, 2009). Text can be presented in either spoken or written form, whereas visualizations include graphs, maps, sketches, photos, videos, or animations. This means that also textbooks, which include pictures can be classified as multimedia learning material, even though only one sensory modality (in this case the visual senses) and only one medium (book) is involved. Throughout this thesis, this broad definition suggested by Mayer will be used.

Overall, adding visualizations to text can be regarded as a promising and successful way to enhance learning. This view is supported by strong empirical evidence that people learn better with multimedia instruction than with text alone as indicated by the reviews of Anglin, Vaez, and Cunningham (2004), Carney and Levin (2002), Fletcher and Tobias (2005), Levie and Lentz (1982), Levin, Anglin, and Carney (1987), as well as Mayer (2001, 2009). However, it should be noted that there are some boundary conditions for this to be true, such as that visualizations should not be implemented when subjects are learning to read, as in this case the visualizations might distract learners from the primary task (cf. Levie & Lentz, 1982). Another boundary condition is that the visualizations must not serve decorational purposes, but need refer to the content of the text (cf. Levie & Lentz, 1982; Levin et al., 1987) – which is the case for the visualizations used in the current study. The finding that people learn better with multimedia instruction than with text alone, which is also called the multimedia effect (Mayer, 2001, 2009), can be considered as the basis for all further research in multimedia learning. If, in general, learning with multimedia would not be better than learning with text alone, there would be no need to examine which combination of text and visualizations would be best, as text alone would be sufficient. Hence, when developing new multimedia instruction, it appears reasonable to first determine – before investigating more sophisticated research questions – whether adding visualizations to text is indeed beneficial for achieving a better understanding of the new material. In the following, a closer look is taken at the reasons for why adding text to visualizations should be beneficial.

2.1 Why Learning with Multimedia Should be Beneficial

There are at least two perspectives to explain why learning with multimedia as compared to text alone should be beneficial (cf. Schmidt-Weigand & Scheiter, 2011). The first view is outcome-oriented in that it focuses on the mental representations people build when learning with multimedia instruction (Chapter 2.1.1). The second perspective is more functional or process-oriented and focuses on the functions text and pictures might play for the cognitive processing of the information that is conveyed through these external representations (Chapter 2.1.2).

2.1.1 The Outcome-Oriented View on Learning with Multimedia

The Cognitive Theory of Multimedia Learning (CTML) of Richard E. Mayer (2001, 2005a, 2009) is the most prominent outcome-oriented view to explain the superiority of learning with text and visualizations as opposed to only text (i.e., multimedia effect). Within the CTML, the multimedia effect is explained by assuming that additional and better developed internal representations result from learning with text and pictures as compared to only text.

As Figure 2.1 illustrates, the CTML postulates different stages of processing information – which can be traced back to three important cognitive theories/models that influenced the CTML, namely the model of different memory stores by Atkinson and Shiffrin (1968), Baddeley's working memory model (1992), as well as Paivio's Dual Coding Theory (1986, 1991).

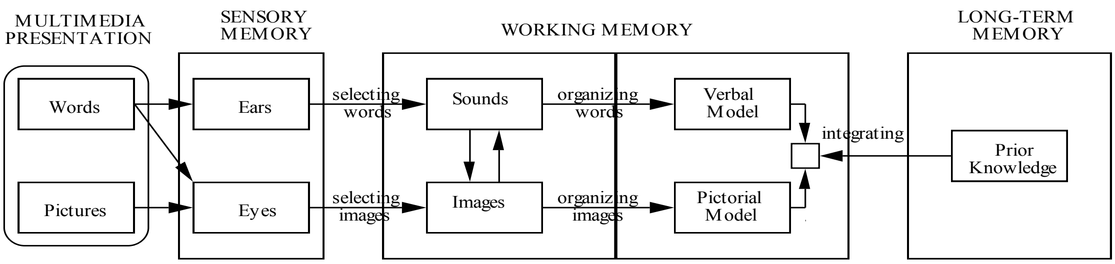


Figure 2.1. The Cognitive Theory of Multimedia Learning (Mayer, 2005a, p. 37).

According to the CTML, multimedia instruction is first processed in the sensory memory before it enters working memory. More precisely, pictures enter sensory memory through the eyes, whereas words enter sensory memory either through the eyes (in case of written text), or through the ears (in case of spoken text).

To enter working memory, the most relevant information from sensory memory needs to be *selected*. It is assumed that in working memory, selected spoken words are initially processed as sound images, while selected pictures are initially processed as visual images. Selected written words are supposed to be initially processed as visual images that are then converted in sound images.

In a next step, through the processes of *organizing* the selected information, the sound images of words are transformed into a verbal mental model, whereas the visual images of pictures are transformed into a pictorial mental model. According to the CTML, sounds and spoken words are processed in an auditory/verbal channel (upper half of Figure 2.1), while pictures are processed in a visual/pictorial channel (lower half of Figure 2.1). Written words, however, are supposed to be initially processed in the visual/pictorial channel according to the sensory modality, and then to be processed in the auditory/verbal channel according to the codality. Thus, verbal information presented in written form will not end up in the pictorial model, but in the verbal model through converting printed words in sounds, given the information is selected and organized appropriately.

The assumption that both, a verbal mental model as well as a pictorial mental model, are constructed is derived from Paivio's Dual Coding Theory (1986, 1991). According to the Dual Coding Theory, the processing of verbal information leads to internal linguistic representations (logogens), while the processing of nonverbal information ends in analogical representations (imagens). At this, processing text and pictures is assumed to result in dual coding, that is in the construction of both, logogens and imagens, whereas processing only text is assumed to result most likely in one code only, namely logogens. Information that is coded in both ways is assumed to be better accessible in memory. Between logogens and imagens referential connections can be established. Thereby, information from one internal representation can be activated by the other internal representation.

In a last step, the verbal and pictorial models are, along with prior knowledge, *integrated* into an integrated mental model. To do so, connections between the verbal and pictorial model are drawn by mapping elements and their relations from one model to the other. Note that also this integration process is somewhat similar to assumption in the Dual Coding Theory (Paivio, 1986, 1991) concerning the building of referential connections between imagens and corresponding logogens (cf. Scheiter, 2009). Prior knowledge that is activated from long-term memory is assumed to be used to coordinate the integration process. As a result, an integrated mental model is formed that, contrary to the two mental models built earlier, is not linked to a specific representational code in the CTML. Rather, it is left open whether the integrated model is either represented in an abstract format, or in a multimodal format. According to the CTML, a deeper understanding of the content can be achieved only if the materials are actively processed, that is if the presented information is selected, organized and integrated. However, active processing requires cognitive resources. In line with contemporary cognitive psychology research, it is assumed that resources, both in sensory memory and working memory are limited. Hence, a major aim when designing multimedia instruction is to help learners in making optimal use of these limited resources.

The idea of the CTML, namely that different internal representations – in this case different mental models – derive from different external representations (such as text and pictures) builds the basis to call this view an outcome-oriented one. The theoretical rationale within the CTML to explain the multimedia effect is based on the mental models resulting from learning with multimedia: When learners are presented with text and visualizations, they are able to build both, a verbal and a pictorial model, and make connections between these two models. Hence, building a verbal and a pictorial model also increases the chances that an integrated mental model will be constructed, and, thus, a deeper understanding of the content is achieved. In contrast, if only text is presented, learners might be able to build a verbal model, but less likely a pictorial model. Thus, they less likely will make connections between these models, resulting in a less developed integrated mental model.

In compliance with this line of reasoning, the benefits of multimedia instruction (i.e., presenting text and visualizations) should be observed most likely for tasks that require predominantly either the use of the pictorial mental model or the integrated mental model, but not necessary for tasks that solely require the use of the verbal mental model.

While there exists no simple one-to-one mapping between different task formats and the knowledge representations they address, some assumptions can be made concerning this mapping. These assumptions are based on two arguments: First retention/recall tasks, which refer directly to what has been stated in the learning materials are more likely to address knowledge stored in the verbal and pictorial mental model, respectively, whereas transfer tasks, which require a deeper understanding of the content are more likely to be based on the integrated mental model. Second, there is some correspondence between the representation format of the task (i.e., verbal vs. pictorial) and the mental model it addresses, in that answers to verbal tasks are more likely to be based on the verbal mental model, while answers to pictorial tasks are more likely to be based on the pictorial mental model. This is assumed to be the case, because under these conditions, no further translation from task format to knowledge format is required when giving an answer; rather, the information can be directly read off the mental representation (cf. Scheiter, Wiebe, & Holsanova, 2008). Despite these correspondences, it certainly depends on what a question asks for, whether the verbal, or the pictorial mental model, or both contribute to the answers and to what extent. For transfer tasks, the representational code of the task should be irrelevant, as the integrated mental model is no longer linked to a specific representational code.

To conclude, to *directly* assess the content of the verbal mental model, verbal factual knowledge tasks may be used that address what was explicitly conveyed through the text. This is most often done in current research on multimedia learning by means of verbal retention tasks, verbal multiple-choice tasks, or cloze texts. To *directly* assess the pictorial mental model explicitly, factual knowledge tasks posed in a pictorial format may be used that address what was explicitly depicted in the visualizations. This can be done, for instance, by administering drawing tasks, picture sorting tasks, or picture identification tasks (e.g., Joseph & Dwyer, 1984). It should be noted though that in current multimedia research the role of the format of the tasks in factual knowledge tasks, and specifically the role of pictorial tasks, is often mostly neglected (cf. Anglin et al., 2004; Scheiter et al., 2008), even though there are some exceptions (e.g., Bartholomé & Bromme, 2009; Brünken, Steinbacher, Schnotz, & Leutner, 2001; Schmidt-Weigand & Scheiter, 2011; Schmidt-Weigand, Kohnert, & Glowalla, 2010).

As mentioned above, the integrated model is constructed by elaborating on the inter-representational connections between the verbal and pictorial mental model, and by activating prior knowledge; therefore, it goes beyond the information contained in these representations. According to the CTML, the quality of the integrated mental model is supposed to be best reflected by performance in transfer tasks – irrespective of other representational formats – where a learned content has to be applied to new situations, thereby proving that a deeper understanding of the content has been acquired (cf. Mayer, 2001, 2005a, 2009).

To sum up, learning with text and visualizations as opposed to only text should result in a more elaborated pictorial as well as in a more elaborated integrated mental model, but not necessarily in a better developed verbal mental model. Hence, in turn, the multimedia effect should be more pronounced for pictorial and transfer tasks as compared to verbal factual knowledge tasks. However, this should not to be misunderstood in a way that the multimedia effect is assumed to not hold true for verbal factual knowledge tasks, but solely that it might be less accentuated for verbal factual knowledge tasks as compared to pictorial factual knowledge tasks or transfer tasks.

In line with this reasoning, the superiority of learning with text and visualizations compared to text-only has been shown to be more pronounced for transfer rather than (verbal) factual knowledge tasks. Mayer (2001) reported in an overview of his own studies that in six of nine studies students achieved higher learning outcomes in verbal factual knowledge tasks; the median of the nine effect sizes was .67, whereas in nine of nine studies, students showed better performance in transfer tests when learning with text and visualizations as opposed to text alone;

the median of the nine effect sizes was 1.50. Hence, larger effect sizes in favor of the multimedia effect were observed for transfer tasks than for verbal factual knowledge tasks. Moreover, with respect to factual knowledge tasks, the superiority of text and visualizations over text alone has been shown to be especially accentuated for pictorial tasks like drawing tasks or picture identification tasks, but less accentuated, and sometimes even nonexistent for verbal tasks like terminology tasks or cloze texts (e.g., Alesandrini & Rigney, 1981; Baker & Dwyer, 2000; Beagles-Roos & Gat, 1983; Joseph & Dwyer, 1984; Szabo, Dwyer, & DeMelo, 1981; for an overview see Levie & Lentz, 1982).

Despite the fact that the outcome-oriented view inherent to the CTML allows predicting multimedia effects for different types of knowledge assessments, nevertheless it also has some limitations. In particular, the CTML does not describe the functions that text and visualizations might play with regard to the cognitive processes they may facilitate, and the properties these external representations might possess to be beneficial for learning. Thus, this topic will be addressed next.

2.1.2 Functional View on Learning with Multimedia

In addition to the abovementioned outcome-oriented view, the superiority of learning with text and visualizations over text alone might be also explained by the cognitive processes that are facilitated when learning with multimedia instruction. At this, text and pictures can facilitate cognitive processing, and thereby foster deeper understanding by 1) computational offloading, 2) re-representation, 3) graphical constraining, and 4) supporting elaborations. These four arguments will be explicated below.

Computational offloading

According to Scaife and Rogers (1996), computational offloading "refers to the extent to which differential external representations reduce the amount of cognitive effort required to solve informationally equivalent problems" (p. 188). For instance, for visuo-spatial information, pictorial in comparison to verbal representations can be more computationally efficient by reducing the need to search, recognize, and memorize information. That is, visual elements representing conceptually related information are grouped together in a pictorial representation, and, hence, this information can directly be read-off (Larkin & Simon, 1987), whereas in text, this information may be distributed across different paragraphs.

Similarly, Levie and Lentz (1982) argued that when visualizations are presented with text, the visualizations can offload working memory as they "may help to keep the relationships between key concepts at ready access, freeing the learner's processing capacity for other aspects of the learning task" (p. 222). Taken together, it can be argued that adding visualizations to text may reduce unnecessary demands imposed on working memory.

Such an argumentation may be well in line with assumptions made by one of the most prominent instructional design theories that focus on the demands on working memory when processing information, namely the Cognitive Load Theory (CLT; Sweller, 1999, 2005a; Sweller van Merriënboer, & Paas, 1998). The CLT is not a theory specifically devoted to multimedia learning, but has been applied to a variety of instructional materials including, among others, multimedia instruction.

According to the CLT, three kinds of cognitive load can be distinguished that constitute the overall load on working memory's limited capacity, namely intrinsic cognitive load (ICL), extraneous cognitive load (ECL), and germane cognitive load (GCL).

ICL is conceptualized as the load on working memory that depends on the element interactivity (or complexity) of the learning material, as well as the expertise (or prior knowledge) of the learner. The higher the number of elements and their interrelations that have to be hold in working memory, the higher is ICL. ICL is reduced once learners have available prior knowledge that helps them to group a number of elements into a larger meaningful unit.

ECL is said to be an unnecessary load on working memory resulting from a bad instructional design, where learners have to conduct cognitive processes that do not contribute to learning. One way to explain the multimedia effect within CLT is that ECL is decreased through adding visualizations to text, since (as explicated above) thereby unnecessary demands on working memory might be reduced. According to the CLT, the main goal of instructional design lies in the reduction of ECL to free working memory resources for more valuable processing activities.

These valuable processing activities that learners have to conduct to foster meaningful learning constitute GCL. Like ECL, the investment of GCL can be altered by instructional design; a further goal of instructional design should hence be to increase a learner’s GCL. How the construct of GCL can also be used to explain the multimedia effect will be explained below (cf. supporting elaborations).

Re-representation

Re-representation “refers to how different external representations that have the same abstract structure, make problem-solving easier or more difficult” (Scaife & Rogers, 1996, p. 189). At this, verbal and pictorial representations may differ in the processes each representation supports, even though the representations contain the same information (cf. Zhang & Norman, 1994). According to Schnotz (2002, 2005), verbal representations are more powerful in conveying abstract knowledge and concepts, which cannot be depicted in a single picture, as pictures refer to concrete concepts. For instance, when trying to convey that food is not allowed in public transportation, one can easily refer to the concept of food by means of a verbal representation. However with pictorial representations, one can refer only to concrete objects (e.g., ice-cream, hamburgers et cetera), but not to an abstract concept. On the other hand, pictorial representations may be more apt in conveying visuo-spatial information. At this, Levie and Lentz (1982) suggested that particularly for visuo-spatial information, visualizations may be an efficient substitute for words.

Graphical constraining

Pictorial representations are often less ambiguous than verbal representations, and hence constrain the interpretation of a representation. For instance, when reading “the bottle is next to the glass” it is unclear whether the bottle is on the left side of the glass or on the right side. In contrast, this relationship is explicitly depicted within a visualization. Thus, pictorial representations are able to reduce the number of possible represented worlds of a topic (cf. graphical constraining, Scaife & Rogers, 1996). The combination of visualizations and text can hence reduce the ambiguity of the text content, as the visualizations may constrain the interpretation of a text segment (Ainsworth, 1999, 2006). Moreover, visualizations may also provide a context in which the textual information can more accurately be interpreted, organized, and understood (Bransford & Johnson, 1972; cf. Levie & Lentz, 1982). Thereby, visualizations may also serve as a source to check one's own understanding of the text (cf. Ainsworth, 1999, 2006; Levie & Lentz, 1982). Similarly, presenting text and visualizations can also reduce uncertainty about the visualizations, as the text might help to better understand the elements of the visualization. Accordingly, when text is presented with visualizations, this might not only offload working memory, but also might lead to less uncertainty regarding the content, which may be reflected in fewer erroneous statements made by learners (Butcher, 2006), fewer negative monitoring statements (i.e., expressions referring to not understanding), as well as more positive monitoring statements (i.e., expressions of understanding) compared to when only text is presented.

Supporting elaborations

By substituting for more demanding reasoning processes, pictorial representations may allow for drawing inferences grounded in perception (Goldstone & Son, 2005; Schwartz, 1995). Several think-aloud studies are in line with this view, as they show that learners conducted more valuable processing activities when processing visualizations than when processing text (Ainsworth & Loizou, 2003; Cromley et al., 2010; Moore & Scevak, 1997). By combining text with visualizations as opposed to only text, a more thorough processing of the content might hence be supported. Moreover, presenting text and visualizations offers the opportunity to relate these different representations to each other (cf. Ainsworth, 1999, 2006). By doing so, a learner may gain insights that he/she could hardly achieve with solely one representation. Also, Levie and Lentz (1982) suggested that visualizations may lead readers to increase their depth of semantic analyses, for instance, by inducing a deeper processing of the text (cf. Peeck, 1993). In line with these arguments, Butcher (2006) observed, by applying the think-aloud methodology that learners generated more inferences when studying text and visualization rather than only text. In terms of CLT (Sweller, 1999, 2005a; Sweller et al., 1998), conducting such valuable processing activities might be interpreted as an increase in GCL.

Empirical evidence

To sum up, based on a functional view, one might gain more insights in the cognitive processes evoked when learning with text and visualizations as opposed to text only. With regard to the processes, which may occur, learning with text and visualizations should lead to less uncertainty about the content, and, furthermore, to a more elaborate processing of the content as compared to learning with text alone.

The empirical support for these assumptions is rather sparse, since these assumptions have seldom been investigated so far. In a study by Butcher (2006), learners receiving text and visualizations made fewer erroneous statements, which might be regarded as an indicator of being more certain about the content. Also, Butcher (2006) observed that learners receiving text and visualizations generated more inferences than learners receiving only text. Similarly, in a recent study by Cromley, Snyder-Hogan, and Luciw-Dubas (2010), the authors observed comparatively more inferences when learners worked with visualizations than when they studied text only. Moreover, this higher proportion of inferences in the visualization condition was related to better performance (see also Ainsworth & Loizou, 2003).

With respect to cognitive load, one might expect that ECL will be decreased, and GCL will be increased when learners are receiving text and visualizations as opposed to learners who are receiving text only. In line with this reasoning, recently Schmidt-Weigand and Scheiter (2011) could show a reduction of ECL when learners received text and visualizations as opposed to learners who received only text. No differences occurred with respect to the item supposed to measure GCL, which might however also be attributed to problems in measuring GCL with the item the authors used in their study (cf. de Jong, 2010).

2.1.3 Conclusions

There is strong empirical evidence that adding visualizations to text is a successful and promising way to foster learning. There are at least two views that might account for why learning with multimedia should be beneficial: The outcome-oriented view stresses the cognitive outcomes achieved through multimedia learning, whereas the functional view emphasizes the cognitive processes that are supported when studying text-picture combinations.

The most prominent outcome-oriented view in multimedia learning, the CTML (Mayer, 2001, 2005a, 2009), suggests that different mental models are built when learning with multimedia. These additional and/or better developed mental models are assumed to result in a better understanding of the content. More precisely, the CTML assumes that through adding visualizations to text, learners will be more likely to additionally build a pictorial mental model. Furthermore, this pictorial mental model might help to build a better developed integrated mental model.

Whereas the outcome-oriented view focuses on the different mental models build when learning with multimedia, the functional view on the other hand focuses on the cognitive processes that might be facilitated when visualizations are added to text. According to the functional view, adding visualizations to text might reduce unnecessary processing demands, which in terms of CLT correspond to a decrease in ECL. Moreover, it might reduce uncertainty about the conveyed content and enable valuable processing activities such as inferences – which in terms of CLT correspond to an increase in GCL.

In the current thesis both approaches are considered and investigated in Study 1. It should be noted that even though it may be argued that under certain circumstances these two approaches lead to differing predictions (cf. Schmidt-Weigand & Scheiter, 2011), for the scope of Study 1 of this thesis, the two approaches do not contradict each other. While assumptions that can be derived from the outcome-oriented view are investigated indirectly by means of assessing performance on different learning outcome measures, the assumptions that can be derived from the functional view are investigated through assessing experienced cognitive load as well as think-aloud protocols. Moreover, these data will be related to each other, to investigate, for instance, whether the number of generated inferences (observed from the think-aloud protocols) is associated with a deeper understanding as measured in transfer tasks, which are an indicator of the integrated mental model.

Overall, the considerations made so far strongly recommend adding visualizations to text. However, these considerations do not provide any conclusion concerning differential effects of learning with different types of visualizations, such as dynamic as opposed to static visualizations.

Since for the domain of the current thesis dynamic visualizations might be even more apt than static visualizations, the topic of learning with these two types of visualizations will be addressed next.

2.2 Learning with Dynamic and Static Visualizations

The domain of the current thesis, the physical principles underlying fish locomotion, is a concrete example that illustrates Newton's laws of motion. This domain comprises various aspects of the fish locomotion, such as the interplay of the changing velocity of a fish's caudal fin along its trajectory, and its impact on the sizes of the associated resulting forces, or its impact on the related swimming speed. One may wonder if for this domain specifically dynamic visualizations as compared to static visualizations might foster learning. This might be the case, because dynamic visualizations seem to possess enormous potential to improve the understanding of such interrelations that change over time (such as the changes in velocity of the caudal fin and its impact on resulting forces).

In the following, first the characteristics of dynamic and static visualizations will be clarified. Thereafter, as a consequence of their properties, the processing demands of dynamic and static visualizations will be analyzed. Finally, an overview of research on learning with dynamic and static visualizations will be given.

2.2.1 Definition of Dynamic and Static Visualizations

As was explicated above, visualizations hold the potential to convey visuo-spatial information in a computational efficient way (cf. Larkin & Simon, 1987). However, it might be the case that for certain learning objectives not only the visuo-spatial information about entities is important, but how these entities change over time. To depict such changes, dynamic visualizations (like animations or videos) might be specifically helpful (cf. Lowe, 2003).

In the following, it will be clarified what can be regarded as dynamic visualizations in contrast to static visualizations. At this, dynamic and static visualizations can be defined according to two aspects: One aspect concerns how dynamic and static visualizations differ from a technical point of view, and the other aspect concerns the information that can be directly depicted. As suggested by Schnotz and Lowe (2003), such differences of dynamic and static visualizations might influence how these two types of visualizations are perceptually and cognitively processed, which will be explicated in more detail in Chapter 2.2.2.

From a technical point of view, a dynamic visualization consists of a series of still frames, which are shown rapidly one after the other, as is suggested by Bétrancourt and Tversky (2000), who state that "computer animation refers to any application which generates a series of frames, so that each frame appears as an alteration of the previous one and where the sequence of frames is determined either by the designer or the user" (p. 313). The number of frames per second constitutes what Schnotz and Lowe (2008) call the temporal granularity, and it influences the perception of a continuous movement. According to Schnotz and Lowe (2008), a dynamic visualization should consist of at least five frames per second to be perceived as continuous. It should be noted that in addition to the number of frames per second, the alteration in space between each single picture affects the impression of a continuous movement (Schnotz & Lowe, 2008; see also Rieber & Kini, 1991). If the temporal granularity was further reduced, a movement would no longer be perceived as continuous, but as a sequence of static frames, where one frame replaces the previous frame. Such a sequence of static frames can be regarded as a presentation format of static visualizations, namely as *static-sequential visualizations*. Another presentation format of static-visualizations that consist of more than one picture, are static-simultaneous visualizations. In *static-simultaneous visualizations*, all pictures are presented next to each other at the same time. In principle, instructional materials involving static visualizations do not necessarily need to consist of multiple static pictures, but they can also comprise a single static picture only. However, when comparing dynamic and static visualizations, it is often advocated to implement multiple static frames for the static visualization condition to keep the depicted information in the different visualization formats as informationally equivalent as possible to have a "fair" comparison (e.g., Tversky et al., 2002). It should be noted that even when using multiple static visualizations, they are not necessarily completely informational equivalent to dynamic visualizations, since these types of visualizations will nevertheless differ with regard to their inherent properties (see below).

In addition to this technical point of view, one can distinguish dynamic and static visualizations with respect to the information that can be directly depicted, that is with regard to their content. According to Rieber and Kini (1991), dynamic visualizations present the changes of the position of an object over time (motion), and also the direction of these changes (trajectory). Additionally, and in contrast to static visualizations, a characteristic of a dynamic visualization is that it "triggers the perception of a continuous change" (Schnotz & Lowe, 2008, p. 304). Concerning the continuous changes that can be displayed in dynamic visualizations, Lowe (2003, 2004) distinguishes three main types of changes: translations, transformations, and transitions. *Translations* refer to position changes of objects from one position to another and correspond to Rieber's and Kini's (1991) concept of motion. *Transformations* refer to changes in an object's appearance, such as changes in size, shape, or color. *Transitions* involve the appearance or disappearance of entities, irrespective of whether an entity is added/removed, or if the entity is moving in/out of the visual field. It may be argued that depending on the type of change that is depicted, dynamic visualizations might be more or less helpful, which is however not in the scope of the current thesis, as all types of changes are relevant to understanding Newton's law of motion in the context of fish locomotion. Translations, transformations, and transitions can be considered as *visuo-spatial changes*. Moreover, dynamic visualizations – in contrast to static visualizations – can not only depict continuously visuo-spatial changes, but also possess the property to directly depict *temporal information*: For instance, dynamic visualizations allow depicting how long it takes for an object to change its position from point A to point B, and if this change is constant (cf. Lowe, 2003; Schnotz & Lowe, 2008). Moreover, dynamic visualizations can exclusively display dynamic features, such as changes in velocity.

These differences between dynamic and static visualizations, namely that dynamic visualizations can continuously display visuo-spatial changes, as well as temporal information (e.g., changes in velocity), in turn may affect how these visualizations are perceptually and cognitively processed. This topic will be discussed next in more detail.

2.2.2 Processing Demands of Dynamic and Multiple Static Visualizations

In the following, as a consequence of their different characteristics, the benefits and challenges of processing dynamic and static visualization formats will be analyzed with respect to the processing demands these types of visualizations impose onto learners (cf. Gerjets, Imhof, Kühl, Pfeiffer, Scheiter, & Gemballa, 2010). At this, it will be distinguished between dynamic visualizations, and two presentation formats of static visualizations: static visualizations where the static pictures are presented sequentially one after the other at the same position (static-sequential), and static visualizations where the static pictures are presented simultaneously next to each other at the same time (static-simultaneous). The processing demands that might be relevant to describe learning from these different visualization formats are related to five aspects, namely 1) access to information concerning visuo-spatial changes, 2) access to information concerning temporal changes, 3) transience, 4) visual complexity, and 5) illusions of understanding. When discussing these processing demands, it will be distinguished between the perceptual and the cognitive processing of these types of visualizations, as suggested by several authors (e.g., Schnotz & Lowe, 2008).

Access to information concerning visuo-spatial changes

Dynamic visualizations present the changes of the position, the form, and the appearance of an entity over time, as well as its direction (cf. Lowe, 2003, 2004; Rieber & Kini, 1991). As dynamic visualizations are able to *continuously* display these visuo-spatial changes of an entity, learners can directly perceive these changes, rather than having to infer them (cf. spatial inferences, Schnotz & Lowe, 2008). In contrast, with static visualizations, learners have to conduct such spatial inferences about the changes of the position, the form, the appearance, and the direction of an entity by means of mental animation (e.g., Hegarty, 1992; Hegarty & Sims, 1994). However, mental animation is supposed to be resource-intensive, and, furthermore harbors the risk that learners might inadequately reconstruct the changes of an entity, which may in turn result in an incomplete or erroneous mental model of the content (Schnotz & Lowe, 2008). Thus, with respect to the demands of mental animation, one can assume that dynamic visualizations might act as an external substitute for the internal processes (cf. supplantation, Salomon, 1979), thereby allowing for cognitive offloading (Scaife & Rogers, 1996), and reducing unnecessary processing demands, namely ECL. This is what Schnotz and Rasch (2005) call the facilitating function of dynamic visualizations. Hence, one can assume that the demands of mental animation might be generally lower for dynamic than any presentation format of static visualizations.

With respect to the different presentation formats of static visualizations, one might argue that the demands of mental animation may be less pronounced for static-sequential as opposed to static-simultaneous visualizations. This might be the case, because spatial changes may be easier retraced, since in static-sequential visualizations these changes are visually superimposed as the new picture appears at the same position as the previous one, so that they might give an impression of (a fragmented) motion (cf. Imhof, Scheiter, & Gerjets, 2009)[1]. In a similar way, Wells, van Mondfrans, Postlethwait, and Butler (1973) explained their finding of a superiority of static-sequential as opposed to static-simultaneous visualizations for conveying concepts of motion. They argued that when a learner viewed a static-sequential visualization, "the subject focused his eyes on one point while the slides changed, making the object's change of position in each succeeding slide more apparent" (p. 239). Hence, static-sequential visualizations might impose less unnecessary demands on working memory (i.e., less ECL) than static-simultaneous visualizations.

Access to information concerning temporal changes

As mentioned above, in contrast to static visualizations, dynamic visualizations do not solely contain information concerning *visuo-spatial* changes, but also possess a *temporal* dimension (cf. Schnotz & Lowe, 2008). Therefore, they can be used to directly depict temporal information. At this, they can depict dynamic features like velocity or acceleration, so that this information can be directly read-off from the visualization. To clarify this point, this will be exemplified in a different domain, namely Kepler's second law. According to Kepler's second law, a planet changes its velocity when orbiting the sun on an ellipse as a function of its distance to the sun. More precisely, the planet is moving faster the nearer it is to the sun, and slower the farer it is away from the sun. This dynamic feature, the change of a planet's velocity, can be directly depicted in a dynamic visualization in a continuous way. In contrast, these dynamic features – in the case of Kepler's second law, the continuous changes in velocity of a planet orbiting the sun – are not inherent properties of static visualizations, and cannot be directly depicted in static visualizations, irrespective of whether the static visualizations are presented sequentially or simultaneously. Moreover, usually such dynamic features cannot even be directly inferred from static visualizations, but have to be given by an additional source (e.g., text or table). Aligning different sources and integrating them mentally and/or conducting inferences about dynamic features (cf.

[1] It should be noted that as with static-sequential visualizations the motion is not displayed continuously, a learner still has to mentally animate the gaps between successive static frames. Hence, also static-sequential visualizations are still supposed to impose processing demands onto a learner due to the need of mental animation.

temporal inferences, Schnotz & Lowe, 2008) might be cognitive highly demanding, leading to an increase in ECL. Lowe (2004, p. 258) arrives at similar conclusions when stating:

> "At best, static depictions can present implicit representations only of dynamic content. They therefore require learners to infer the situational dynamics. […] In contrast, animations have the advantage of being able to present the situational dynamics explicitly and appropriately so that the majority of learners' processing capacity could be devoted to comprehending the content directly."

Moreover, the process of inferring the dynamics in static visualizations can also be an error-prone process (Schnotz & Lowe, 2008). To conclude, dynamic visualizations might be best suited to convey dynamic features like changes in velocity.

Transience

In dynamic visualizations the presented information changes permanently. As a consequence, when learning from dynamic visualizations, previously presented information has to be kept in working memory, while new incoming information has to be processed and then integrated with the information held active in working memory. This may cause a *temporal* split-attention effect (cf. van Gog, Paas, Marcus, Ayres, & Sweller, 2009), and may impose high cognitive demands on working memory (e.g., Ayres & Paas, 2007). However, if the changes occur in the same way several times (i.e., repetitive changes), for instance, in the domain of planetary motion, this possible negative effect of dynamic visualizations should be reduced (cf. Schnotz & Lowe, 2008). This also accounts for the domain of the studies of the current thesis: The movement that has to be understood in the studies of the current thesis, namely the movement of an undulatory swimming fish repeats itself regularly in a periodical way, and is shown several times[2].

With respect to static visualizations, it should be noted that also with static-sequential visualizations the information is virtually transient, because each new static picture replaces the previous one. Also for static-sequential visualizations this possible negative effect should be reduced if the changes are shown several times. For static-simultaneous visualizations, problems associated with transience should not occur at all, and hence should not yield higher working memory demands, since the information conveyed through all static pictures remains on the screen and can directly be read-off. Thereby, with static-simultaneous visualizations, direct visual comparisons between different steps in a sequence can be made, and, moreover, the learner can regulate his own pace of processing the visualizations (cf. Boucheix & Schneider, 2009; Hegarty, Kriz, & Cate, 2003; Mayer, Hegarty, Mayer, & Campbell, 2005).

[2] Note that it is important that not only the movement is repetitive, but also that this movement is shown several times.

Visual complexity

Due to the continuous movement that is exclusively displayed in dynamic visualizations, it might be perceptually more demanding to select the relevant information (e.g., Jarodzka, Scheiter, Gerjets, & van Gog, 2010). This problem is getting even worse, since in dynamic visualizations several elements may change simultaneously at different locations at the same time. Thereby, on the one hand, thematically less relevant information might distract attention away from the thematically more relevant information (Lowe, 1999). Moreover, due to the movement of several elements at different locations at the same time, learners are forced to *spatially* split their attention (e.g., Ayres & Sweller, 2005). Lowe (2003) refers to this as an intra-representation split-attention effect occurring within visualizations, as opposed to an inter-representation split-attention effect that may occur when learners have to attend to two different representations (e.g., text and visualizations). These factors (i.e., continuous movement, distracting movement, intra-representational split-attention) might induce a high degree of visual complexity in dynamic visualizations, which in turn may impose high demands on learners (e.g., Ayres & Paas, 2007; Hegarty, 2004; Hegarty & Kriz, 2008; Lowe, 1999, 2003, 2004; Tversky et al., 2002). In contrast, with static visualizations, single relevant states of a changing system can be accentuated. Moreover, as there is no continuous change in static visualizations, identifying relevant information from a static visualization might be perceptually less demanding, so that the external representations may be more likely to be accurately perceived and comprehended (apprehension principle; Tversky et al., 2002).

Concerning different presentation formats of static visualizations, the apprehension principle may be more accentuated for static-sequential as opposed to static-simultaneous visualizations. While in static-sequential visualizations there is always only one picture presented at the same time, in static-simultaneous visualizations, on the other hand, several pictures are presented at once. At this, with static-simultaneous visualizations, learners have to decide when to attend to which information in the different pictures, and to match the different parts of the visualization. These search and match processes may burden the limited capacity of working memory, and may also be perceptually demanding. This problem may be even more pronounced, since in static-sequential visualizations each picture can be shown in a large size on the computer screen, whereas in static-simultaneous visualizations, each single picture usually needs to be shown in a smaller size to fit the screen when presented next to each other. It should be noted that the smaller size of each picture in static-simultaneous visualizations is not inevitable, as, for instance, the size of each picture in static-sequential visualizations could be artificially decreased to “control” this factor, or the static-simultaneous visualizations could be shown on a larger screen. Nevertheless, even though differences in picture size are not inherent properties of the two visualization formats, for practical reasons these differences are present in many instructional materials.

Illusion of understanding

It is sometimes argued that since dynamic visualizations like videos and animations are commonly associated with entertainment – as opposed to any presentation format of static visualizations – they may seem to be easy to understand, which in turn may result in an illusion of understanding (e.g., Bétrancourt, 2005; Lewalter, 2003; Rebetez, Bétrancourt, Sangin, & Dillenbourg, 2010). In a similar way, Schnotz and Lowe (2008) stated that because of their association with entertainment, dynamic visualizations may trigger processing strategies that are inadequate for learning. As a consequence, at worst, learners may disengage from deeper processing of the content (underwhelming, Lowe, 2003, 2004; Salomon, 1984; see also inhibiting function, Schnotz & Rasch, 2005). It should be noted that in most studies in which similar arguments have been made, there has been no direct empirical evidence in favor of an illusion of understanding, or of a shallower processing; rather these assumptions have been generated post-hoc in order to explain the observed results. A possible exception is a study by Lewalter (2003), where learners were asked to think aloud while learning from either dynamic or static visualizations. She observed more positive monitoring statements in the dynamic as opposed to the static visualization condition. At the same time, learners in the dynamic visualization condition did not outperform learners in the static visualization condition. Taken together, this may be interpreted as evidence for an illusion of understanding occurring during learning from dynamic visualizations. On the other hand, it should also be noted that even though an illusion of understanding might have occurred when learning with dynamic as opposed to static visualizations, this did not lead to a shallower processing of the content in Lewalter's study (as measured by means of elaborations).

2.2.3 Conclusions for the Current Studies

To sum up, when considering the processing demands of dynamic as opposed to static visualizations, one potential challenge of dynamic visualizations related to processing may be that they evoke an illusion of understanding, even though this claim has hardly been directly investigated. Another potential challenge of dynamic visualizations might lie in their transient nature, which, however, is considered to be diminished for dynamic visualizations that show a repetitive movement several times. This is the case for the dynamic visualizations of the domain of fish locomotion used in the current studies. A last potential challenge of dynamic visualizations is that they might suffer from a higher visual complexity than any format of static visualizations. This might also account for the dynamic visualizations in the current study, since several elements are changing continuously at the same time, making them particularly perceptually harder to process.

However, specifically for a domain where the visuo-spatial changes might be hard to mentally animate, dynamic visualizations might be especially helpful, because they depict these changes in a direct way. Likewise, if for a given domain the understanding of temporal changes, such as changes in velocity is crucial, dynamic visualizations might be best suited to convey such information. At this, dynamic visualizations, may reduce unnecessary demands on working memory (i.e., decrease ECL), because learners do not need to engage in resource-demanding and error-prone processes like spatial and/or temporal inferences. Since ECL is supposed to be reduced, cognitive resources might be available that can be devoted to more valuable processing activities (i.e., increase GCL).

The domain of this thesis deals with the physical principles underlying undulatory fish locomotion and is characterized on the one hand by visuo-spatial changes that can be regarded as hard to mentally animate, and, furthermore, by temporal changes, namely changes in velocity. More precisely, this domain addresses the interplay of the trajectory and changes in velocity of a fish's different body parts, the changing sizes of the associated resulting forces, and the related swimming speed. Therefore, based on this analysis, and its relation to the chosen kind of multimedia instruction, even though the used dynamic visualizations may be perceptually more demanding, they nevertheless are supposed to be better suited for conveying a deeper understanding of the dynamics underlying the given domain than static visualizations, irrespective of how these static visualizations are presented.

With regard to the presentation format of static visualizations, both, static-sequential and static-simultaneous visualizations, possess advantages and drawbacks. This is the case, as on the one hand, learning with static-simultaneous visualizations may reduce processing demands, since the information that is depicted by the visualizations can directly be read-off. On the other hand, however, static-simultaneous visualizations may load a learner’s working memory, as the content may be harder to mentally animate, and, furthermore, decisions have to be made when and how to attend to which information. Moreover, static-simultaneous presented visualizations might be perceptually more demanding than static-sequentially ones, since more visual search and matching processes have to be conducted. This argument will be taken up in the context of cueing visualizations (Chapter 4.2). It should be noted that in the remainder of this thesis the distinction of static visualizations in static-sequential and static-simultaneous visualizations will be made only if the line of reasoning solely accounts for one of these two types of static visualizations. Otherwise, the broader term static visualizations will be used.

In the following, an overview on the research of learning with dynamic and static visualizations will be given. At this, factors will be considered that were emphasized in recent research and that might influence the effectiveness of learning with these types of visualizations, since these factors are also related to the aforementioned processing demands in different ways.

2.2.4 Overview on Research of Learning with Dynamic and Static Visualizations

In the reminder of this chapter, first, an overview on the current state of research learning with dynamic and static visualizations will be given on a global level. Thereafter, factors will be introduced that were emphasized in more recent research on learning with these types of visualizations.

2.2.4.1 Global comparisons

At first glance, from a naïve point of view, it may seem plausible to assume that learning with dynamic visualizations might be more apt than learning with static visualizations. However, when considering the research literature on learning with dynamic and static visualizations, the picture that arises remains somewhat unclear (e.g., Höffler & Leutner, 2007; Tversky et al., 2002), and will be explicated next.

In a review by Park and Hopkins (1993), the authors examined 27 studies, in which 15 studies demonstrated differences in favor of dynamic visualizations, whereas for the other 12 studies dynamic and static visualizations yielded equal learning outcomes.

Bétrancourt and Tversky (2000) reviewed 10 studies comparing dynamic and static visualizations, of which six showed a superiority of dynamic over static visualizations, whereas for the remaining four studies no differences were observable[3].

In a critical review by Tversky et al. (2002), the authors had a rather discouraging view on learning with dynamic compared to static visualizations. In their analysis, they listed eleven studies where no differences between dynamic and static visualizations could be observed. More importantly, however, they re-examined further eleven selected studies and could trace back the observed effects of a superiority of dynamic over static visualizations to an inequality in either content depicted by the visualizations, or procedures associated with dynamic visualizations. That is, on the one hand, in several studies, static visualizations were not informationally equivalent to dynamic visualizations. In these studies a lack of informational equivalence was not only due to inherent properties of the visualizations themselves that possibly cannot be circumvented (e.g., because dynamic visualizations, but not static visualizations, allow showing acceleration). Rather, in these studies the authors had failed to at least approximate informational equivalence by, for instance, comparing dynamic visualizations to only one static picture instead of multiple static pictures. Moreover, besides not being informationally equivalent, in some comparisons of

[3] Note that also additional studies were reviewed that, however, did not focus on a comparison of dynamic and static visualizations, and hence will not be considered here any further.

dynamic and static visualizations the procedures for dealing with the visualizations were not equal (e.g., interactivity was implemented in dynamic visualizations, whereas this was not the case for static visualizations). Note that Tversky et al. (2002) did not rule out the possibility that dynamic visualizations might be superior to static visualizations for specific purposes, such as for depicting continuous changes in time or spatial transformations. However, they suggested that when comparing dynamic and static visualizations, these comparisons should be implemented in a way that differences can be attributed to the type of visualizations per se, and not to unequal information or procedures. It should be noted that even in newer studies (i.e., studies published after the review by Tversky et al., 2002), this asking is by no means self-evident, as can be for instance seen below in Table 2.1, in which from the 34 studies, eleven studies were not in line with this recommendation.

A meta-analysis by Höffler and Leutner (2007) revealed a medium-sized overall advantage of dynamic over static visualizations. In this meta-analysis 26 primary studies were analyzed with 76 pair-wise comparisons of dynamic and static visualizations. One major criterion for the selected studies was that they should correspond to the recommendations of Tversky et al. (2002). That is, that the visualizations were basically informational equivalent, and that studies containing interactive dynamic visualizations were omitted. Of these 76 comparisons, 21 showed a significant superiority of dynamic over static visualizations, two showed a significant superiority of static over dynamic visualizations, and for the remaining 52 comparisons no significant differences were observable. However, even though only in 21 of 76 studies dynamic visualizations were statistically superior, 54 studies at least indicated an advantage of dynamic visualizations on a descriptive level, which in total probably accounted for the medium-sized overall effect of d = .37 .

Since this meta-analysis, which included papers published until the year 2004, further studies were conducted that compared the effectiveness in learning with dynamic and static visualizations. These studies are listed in Table 2.1.

Table 2.1

Effects in Favor of Dynamic Visualizations from Studies Comparing Learning with Dynamic and Static Visualizations Published Since 2004

	Authors	Effect for Dynamic Visualizations
1	Ardac & Akaygun, 2005 [a]	Positive
2	Arguel & Jamet, 2009	Positive
3	Ayres, Marcus, Chan, & Qian, 2009 (Exp. 1)	Positive
4	Ayres, Marcus, Chan, & Qian, 2009 (Exp. 2)	Positive
5	Boucheix & Guignard, 2005	Positive
6	Boucheix & Schneider, 2009 [b]	Positive
7	Fischer, 2008 (Exp. 2) [b]	Positive
8	Höffler, 2007 (Exp. 1)	Positive
9	Höffler, 2007 (Exp. 2)	Positive
10	Imhof et al., 2009 [b]	Positive
11	Imhof, Scheiter, Gerjets, & Edelmann, 2010	Positive
12	Iskander & Curtis, 2005 [a]	Positive
13	Kim, Yoon, Whang, Tversky, & Morrison, 2007 [b]	Positive
14	Kriz & Hegarty, 2007 [a]	Positive
15	Lin, Chen, & Dwyer, 2006	Positive
16	Lin & Dwyer, 2010	Positive
17	Marbach-Ad, Rotbain, & Stavy, 2008 [a]	Positive
18	Münzer, Seufert, & Brünken, 2009 [b]	Positive
19	Pfeiffer, Gemballa, Jarodzka, Scheiter, & Gerjets, 2009	Positive
20	Rebetez et al., 2010	Positive
21	Schnotz & Rasch, 2005 [a]	Positive
22	Stebner, 2009	Positive
23	Wang, Vaughn, & Liu, 2011 [a]	Positive
24	Watson, Butterfield, Curran, & Craig, 2010	Positive
25	Wong et al., 2009 (Exp. 1)	Positive
26	Wong et al., 2009 (Exp. 2)	Positive
27	Wong et al., 2009 (Exp. 3)	Positive
28	Yarden & Yarden, 2010 [a]	Positive
29	Höffler, 2007 (Exp. 3)	Neutral
30	Höffler, Prechtl, & Nerdel, 2010	Neutral
31	Kalyuga, 2008	Neutral
32	Koroghlanian & Klein, 2004	Neutral
33	Tunuguntla et al., 2008 [a]	Neutral
34	van Oostendorp & Beijersbergen, 2007	Neutral
35	van Oostendorp, Beijersbergen, & Solimani, 2008 [a]	Neutral
36	Zhu & Grabowski, 2006 [a]	Neutral
37	Lowe, Schnotz, & Rasch, 2011	Negative
38	Mayer et al., 2005 (Exp. 1) [a]	Negative
39	Mayer et al., 2005 (Exp. 2) [a]	Negative
40	Mayer et al., 2005 (Exp. 3) [a]	Negative
41	Mayer et al., 2005 (Exp. 4) [a]	Negative
42	Scheiter, Gerjets, & Catrambone, 2006	Negative

[a] Note. Studies that are not methodological sound, for instance, in terms of a “fair” comparison of dynamic and static visualizations, as recommended by Tversky et al. (2002).
[b] Note. These studies included either different types of dynamic visualizations or different types of static visualizations, where not every comparison was in favor of dynamic visualizations.

Of these 34 studies with 42 experiments, which are listed in Table 2.1, 28 experiments (66.67 %) found a significant superiority of dynamic over static visualizations for at least one knowledge test

(Chi-square = 4.67; p = .03). Eight experiments (19.05 %) found no differences between dynamic and static visualizations, and six experiments (14.29 %) even found a superiority of static over dynamic visualizations. When considering only studies that conform to the recommendations stated by Tversky et al. (2002) – an equivalent amount of information and procedures in dynamic and static visualizations for at least one comparison – and that can be regarded as methodological sound, fourteen of the 42 experiments need to be excluded (marked in Table 2.1). Of these remaining 28 experiments, 21 experiments (75.00 %) showed a significant superiority of dynamic over static visualizations for at least one knowledge test (Chi-square = 7.00; p < .01), five experiments (17.86 %) showed no differences, and two experiments (7.14 %) showed a superiority of static over dynamic visualizations. When additionally neglecting the five studies in which not every comparison between dynamic and static visualizations was in favor of dynamic visualizations, of the remaining 23 experiments still 16 experiments (69.57 %) showed a superiority of dynamic over static visualizations, even though the comparison only marginally reached statistical significance (Chi-square = 3.25; p = .06). Broadly speaking, these findings mirrored the results of the meta-analysis by Höffler and Leutner (2007): The research on learning with dynamic as opposed to static visualizations is not as discouraging as suggested by Tversky et al. (2002). There might be an advantage of dynamic visualizations, but – as will be discussed in the remainder of this section – whether dynamic visualizations will be superior to static visualizations may depend on specific boundary conditions.

Accordingly, as indicated by the global analysis of the reviews and meta-analysis, it might not be reasonable to question whether learning with dynamic as opposed to static visualizations is more beneficial. Instead of such a global comparison of dynamic and static visualizations, it might be more fruitful to take a more differentiated view into account, and to specify under which conditions this might be the case (e.g., Bétrancourt, 2005; Hegarty, 2004; Plass, Homer, & Hayward, 2009; Schnotz & Lowe, 2008). At this, in recent research several potential moderators were emphasized, namely the *learning objective*, the *presentation format of static visualizations*, *learner characteristics* as well as *design characteristics* (Bétrancourt & Tversky, 2000; Höffler & Leutner, 2007; Park & Hopkins, 1993; Tversky et al., 2002) that – among other things – are also related to the aforementioned processing demands.

2.2.4.2 Recent research concerning potential moderators

Learning objective

When considering a comparison of dynamic and static visualizations, potential benefits of dynamic visualizations may depend on the to-be-achieved learning objective. At this, Bétrancourt and Tversky (2000) recommended to focus on the observed learning outcomes that either require learners to give their answer on the basis of the explicit conveyed content, like in factual knowledge tasks, or on basis of inferences drawn from this explicit conveyed content, like in transfer tasks. According to Bétrancourt and Tversky, differences in learning from dynamic and static visualizations should hardly affect the memorization of explicitly conveyed content, because "explicit knowledge can be retrieved from surface processing structures (such as the text based representations, or the mental model of the picture)" (p. 322). Rather, differences in the learning outcomes achieved by studying dynamic or static visualizations should mainly affect tasks requiring a deeper understanding of the content where it is crucial to successfully draw inferences, as is the case for transfer tasks[4]. Therefore, with respect to the kind of knowledge tasks introduced earlier, namely, verbal factual knowledge tasks, pictorial factual knowledge tasks, and transfer tasks, one would thus mainly expect differences in transfer tasks when learning with either dynamic or static visualizations.

Höffler and Leutner (2007) also examined in their meta-analysis the influence of learning outcome measures, namely declarative knowledge, which corresponds to factual knowledge, and problem solving, which in this case corresponds to transfer knowledge[5]. For both types of knowledge, dynamic visualizations were more apt than static visualizations with effect sizes that can be classified as educational meaningful. However, and surprisingly, this effect was not differently pronounced for these two types of knowledge tasks, which is not in line with the theoretical considerations proposed by Bétrancourt and Tversky (2000). It should be noted though that both factual knowledge tasks and transfer tasks were only assessed in some of the cited studies. Therefore, it cannot be ruled out that this finding can be traced back to the fact that these studies assessing factual knowledge tasks differed from those assessing transfer tasks on further important dimensions, such as the degree to which the applied visualizations were decorational. Generally, the meta-analysis by Höffler and Leutner (2007) should be treated with

[4] It should be noted though that this is only assumed to be the case, if the content to be animated depicts, at a minimum, changes over time, such as motion, so that there is a reasonable justification to use dynamic visualizations at all (cf. Bétrancourt & Tversky, 2000; Park & Hopkins, 1993; Rieber & Kini, 1991).

[5] For reasons of consistency, in the following it will be referred to factual knowledge and transfer knowledge.

caution when aiming at detecting moderating variables due to the rather low number of included studies (26), and pair-wise comparisons (76), respectively. Hence, the assumption that differences in dynamic and static visualizations might be most likely observed for transfer tasks will not be rejected at this point.

To sum up, dynamic visualizations might be best suited for tasks asking for a deeper understanding, such as transfer tasks, given the premise that the content that should be conveyed contains at a minimum changes over time.

Presentation format of static visualizations

When comparing dynamic to static visualizations, recently the influence of design characteristics of the static visualization conditions, specifically the presentation modes of static visualizations, has become a topic of major interest (Boucheix & Schneider, 2009; Imhof et al., 2009; Imhof et al., 2010; Kim et al., 2007; Lowe et al., 2011; Wells et al., 1973). These different presentation modes of static visualizations (i.e., static-sequential and static-simultaneous visualizations) may have an influence on their instructional effectiveness, since – as explicated previously – they may impose different processing demands onto learners.

Up to now, only little research has been conducted regarding the comparison of dynamic visualizations to different formats of static visualizations. These studies yielded inconclusive results: While in some studies a superiority of dynamic visualizations over static-sequential, but not over static-simultaneous visualizations was found (Boucheix & Schneider, 2009; Imhof et al., 2009), in other studies a superiority of dynamic over static-simultaneous, but not over static-sequential visualizations was observed (Kim et al., 2007). Again, other studies found that dynamic visualizations were superior to both, static-sequential and static-simultaneous visualizations (Imhof et al., 2010; Wells et al., 1973). A study by Lowe et al. (2011) even revealed another pattern of results, in that static-sequential visualizations were superior to dynamic as well as static-simultaneous ones.

To roughly summarize these studies: In most cases dynamic visualizations were superior to at least one type of static visualizations, whereas no clear advantage could be shown in favor of any of this two static formats. Nevertheless, to rule out that potential differences between dynamic and static visualizations are valid only for one presentation format of static visualizations, it might be reasonable to compare dynamic visualizations with static-sequential as well as static-simultaneous visualizations to control for this factor[6].

[6] For reasons of economy, this was only done after the instructional material was considered to be optimized, that is, in Study 3 of this thesis.

Learner characteristics

It is often recommended to take learner characteristics into account when considering learning with dynamic and static visualizations (e.g., Bétrancourt & Tversky, 2000; Boucheix & Schneider, 2009; Hegarty & Kriz, 2008; Höffler & Leutner, 2007; Park & Hopkins, 1993). Thereby, in recent research especially the role of *spatial abilities* received attention, since they are reckoned to play an important role in learning with dynamic and static visualizations (cf. Hegarty & Kriz, 2008; Hegarty & Waller, 2005; Höffler, 2010).

The construct of spatial abilities consists of several factors (see Hegarty & Waller, 2005 for an overview). The most prominent distinction is the one by Carroll (1993), which in turn is based on Lohman, Pellegrino, Alderton, and Regian (1987). According to Carroll (1993), spatial abilities consist of five factors, namely *visualization (VZ)*, *spatial relations (SR)*, *closure speed (CS)*, *closure flexibility (CF)*, and *perceptual speed (PS)*. As defined by Carroll (1993), "tests of factor VZ emphasize power in solving increasingly difficult problems involving spatial forms, whereas tests of factor SR emphasize speed in solving relative simple spatial analysis problems" (p. 315). In recent research on learning with dynamic and static visualizations, especially the factor of VZ, and occasionally also SR, are supposed to play an important role in learning with the different visualization formats (e.g., Boucheix & Schneider, 2009; ChanLin, 2000; Hays, 1996; Höffler, 2007; Huk, 2006; Koroghlanian & Klein, 2004; Stebner, 2009; Yang, Andre, & Greenbowe, 2003; see also Höffler, 2010, for a review), whereas the factors CS, CF, and PS are usually neglected. Accordingly, in almost all studies where spatial abilities were assessed, the spatial ability tests used were part of the factor VZ, or SR (Blake, 1977; ChanLin, 2000; Hegarty et al., 2003; Koroghlanian & Klein, 2004; Stebner, 2009), or they consisted of a mixture of several subtests, where at least one test loaded on the factor VZ or SR, respectively (Boucheix & Schneider, 2009; Hays, 1996; Münzer et al., 2009).

In general, there is mostly a positive relationship between spatial abilities and performance when considering learning with dynamic and static visualizations (e.g., Hegarty et al., 2003; Imhof et al., 2009, 2010; Large, Beheshti, Breuleux, & Renaud, 1996; Münzer et al., 2009; Narayanan & Hegarty, 2002; Stebner, 2009; Wender & Mühlböck, 2003; see also Hegarty & Kriz, 2008 as well as Höffler, 2010, for a review), indicating that stronger spatial abilities are beneficial for both, learning with dynamic as well as static visualizations. This may be construed in a way that high spatial abilities allow learners to better perceive and extract visual information in learning with dynamic visualizations on the one hand, and to better mentally animate when learning with static visualizations on the other hand (cf. Hegarty & Kriz, 2008).

Nevertheless, this finding does not rule out that the influence of spatial abilities is differently pronounced when learning with dynamic and static visualizations. According to Mayer and Sims (1994), there are two conflicting hypotheses concerning the moderating role learning prerequisites can play in learning with dynamic visualizations: The ability-as-enhancer hypothesis, and the ability-as-compensator hypothesis. The ability-as-enhancer hypothesis states that if learners possess better learning prerequisites, they will benefit more strongly from dynamic visualizations. Conversely, the ability-as-compensator hypothesis states that especially learners with weaker learning prerequisites will benefit from dynamic visualizations, since they might be overwhelmed by inference and mental animation processes that are required when learning with static visualizations. When considering spatial abilities as a learning prerequisite, the ability-as-compensator hypothesis is mostly advocated. Accordingly, especially learners with weaker spatial abilities might struggle when having to mentally animate changes when receiving static visualizations, so that for these learners dynamic visualizations might be particularly helpful. This is assumed to be the case, because on the one hand, the process of mental animation, which has to be conducted when receiving static visualizations, is highly correlated with spatial ability (e.g., Hegarty & Kozhenikov, 1999; Hegarty & Sims, 1994). On the other hand, when learners with low spatial abilities receive dynamic visualizations, they solely need to perceive the visuo-spatial changes, a process, which might be less dependent from spatial abilities. In line with this reasoning, Hays (1996) observed that learners with low spatial abilities profited more from dynamic than from static visualizations. Moreover, for learners with stronger spatial abilities, whose working memory is not highly loaded when mentally animating, benefits of presenting dynamic visualizations might emerge to a lesser extent. Therefore, one might argue that dynamic visualizations might act as a compensator for learners with weaker spatial abilities, and hence might play a moderating role in learning with dynamic and static visualizations (cf. Hegarty & Kriz, 2008; Höffler, 2010).

There are several pieces of evidence that speak in favor of the ability-as-compensator hypothesis. On the one hand, there are a couple of studies, in which a moderating role of spatial abilities in learning with dynamic and static visualizations could be observed (Blake, 1977; Boucheix & Schneider, 2009; Höffler, 2007; Exp. 1 & Exp. 2), thereby supporting the ability-as-compensator hypothesis. Moreover, the influence of spatial abilities on learning with either dynamic or static visualizations was examined in a recent meta-analysis by Höffler (2010) in which studies were incorporated that investigated the influence of spatial abilities on either static visualizations, or dynamic visualizations, or on both, dynamic and static visualizations. Results revealed that spatial abilities had a positive influence on learning with both, dynamic as well as static visualizations. Furthermore, in line with the ability-as-compensator hypothesis, the mean effect size for the influence of spatial abilities on learning with static visualizations was significant higher than the mean effect size for the influence of spatial abilities on learning with dynamic visualizations.

It should be noted that somewhat related to the ability-as-compensator hypothesis, there is some evidence (Schnotz & Rasch, 2005; Exp. 1) that learners with weaker learning prerequisites (however in this case a combination of intelligence and prior knowledge, but not spatial abilities) learned longer with static visualizations than their counterparts receiving dynamic visualizations. This finding might be interpreted as suggesting that learners provided with static visualizations may possibly try to compensate for the demands of mental animation by watching the static visualizations longer than learners who are provided with dynamic visualizations. Such a compensation strategy might be more accentuated for learners with weaker learning prerequisites.

The abovementioned explanations concerned the influence of spatial abilities in learning with dynamic as opposed to static visualizations in general. However, one may assume that the presentation format of static visualizations might also have implications with regard to the ability-as-compensator hypothesis. More precisely, in Chapter 2.2.2 it was argued that the demands of mentally animating visuo-spatial changes with static-sequential as opposed to static-simultaneous visualizations might be less pronounced. If this would be the case, the assumed moderating role of spatial abilities might be more accentuated, when comparing dynamic visualizations with static-simultaneous visualizations as opposed to static-sequential visualizations. However, it should also be noted that there is no empirical evidence for this assumption in the few conducted studies, which investigated this topic (Boucheix & Schneider, 2009; Imhof et al., 2010).

To sum up, especially when considering learning with dynamic and static visualizations, in line with the ability-as-compensator hypothesis, one might expect dynamic visualizations to be particularly helpful for learners with weaker spatial abilities. Thereby, it is recommended to use a spatial ability test that belongs to the factors VZ or SR, respectively.

Design characteristics of dynamic and static visualizations

In the abovementioned reviews and meta-analysis (Bétrancourt & Tversky, 2000; Höffler & Leutner, 2007; Park & Hopkins, 1993; Tversky et al., 2002), it has been suggested to take design characteristics of the multimedia instruction into account, since they may influence learning with dynamic and static visualizations differently. This suggestion was taken up in several of the more recent studies. At this, as abovementioned, some of these studies focused on the presentation format of static visualizations (e.g., Boucheix & Schneider, 2009; Imhof et al., 2009, 2010; Kim et al., 2007; Lowe et al., 2010). Several of the other studies aimed at counteracting problems in learning with dynamic as compared to static visualizations.

At this, in most of these studies, a goal was to cope with the problem of transience in learning with dynamic visualizations as opposed to static visualizations. For instance, Arguel and Jamet (2009) as well as Rebetez et al. (2010) investigated the influence of adding static pictures to an animation (Rebetez et al. additionally varied learning alone as opposed to learning in collaboration). Whereas adding snapshots to dynamic visualizations was beneficial in the study by Arguel and Jamet (2009; Exp. 1), and led to better learning outcomes as compared to dynamic visualizations without snapshots or static visualizations, in the study by Rebetez et al. (2010), it had no influence on learning outcomes; rather, there was a main effect in favor of dynamic visualizations.

Another potential solution to reduce the problems associated with transience in dynamic visualizations compared to static visualizations might lie in giving learners control over the pacing of dynamic visualizations (i.e., self-pacing). This was explicitly examined in a study by Kriz and Hegarty (2007; Exp. 1) as well as in a study by Boucheix and Guignard (2005). Results in both studies revealed that dynamic visualizations led to better performance compared to static visualizations, while the absence or presence of self-pacing had no effect. Other studies did not directly explore the absence and presence of self-pacing, but solely implemented dynamic visualizations that could be paced by learners and compared them to static visualizations. While in a study by Scheiter et al. (2006), self-paced dynamic visualizations led to even inferior results compared to static visualizations, in a study by Wang et al. (2011), self-paced dynamic visualizations led to a better performance than static visualizations. To sum up, the research on how to cope with the demands of dynamic visualizations associated with transience is inconclusive. Irrespective of that fact, and as was explicated in Chapter 2.2.2, transience is assumed to play a negligible role for the dynamic visualizations used in the studies of the current thesis.

However, transience is not the only potential drawback in learning with dynamic visualizations, since they may also possess a comparatively high degree of visual complexity. But, there are hardly any studies that examined how to improve dynamic visualizations (in comparison to static visualizations) by reducing this latter drawback. This point will be taken up again in Chapter 4, when design characteristic that aim at reducing the processing demands associated with the visual complexity of dynamic visualizations will be described in more detail.

In a nutshell, this literature overview based on the existing reviews, the meta-analysis, and the recently conducted studies point to the fact that the comparison of dynamic and static visualizations should go beyond a global comparison (e.g., Bétrancourt, 2005; Hegarty, 2004; Scheiter & Gerjets, 2010). Thereby, first, it should be assured that it is a fair comparison as recommended by Tversky et al. (2002). When taking a more differentiated view, one might consider why and for which purpose (learning objective), as well as for whom (learner characteristics) one might expect differences when learning with dynamic and static visualizations. Furthermore, it seems reasonable to investigate the potentials of dynamic visualizations as compared to static visualizations by taking design characteristics into account that aim at solving particular problems that are associated with learning from dynamic visualizations. By doing so, more meaningful conclusions might be drawn about the instructional effectiveness of dynamic as opposed to static visualizations.

2.3 Conclusions

Taken together, in line with the CTML, one would assume a better developed pictorial and integrated mental model when learning with multimedia as compared to text alone. Since the pictorial and integrated mental model might be best reflected by pictorial and transfer tasks, respectively, the benefits of multimedia should particularly emerge for these tasks. Derived from the functions that visualizations might play when added to text, learning with multimedia should, on the one hand, reduce unnecessary processing demands (i.e., ECL), and, moreover, should support a more elaborate processing of the content (i.e., GCL).

With regard to different types of visualizations, for the domain at hand – a domain where the understanding of the interrelations of dynamic features is crucial – one might consider dynamic visualizations, as compared to static visualizations, to be especially helpful. This is assumed to be the case, because on the one hand, a potential drawback of dynamic visualizations, namely their transience is diminished, since the depicted movement is shown several times. Moreover, the dynamic visualizations for the domain at hand allow for an immediate and continuously access to information concerning visuo-spatial changes, as well as temporal changes. Thereby, in contrast to learning with static visualizations, learners do not need to conduct resource intensive and error-prone processes like mentally animating visuo-spatial changes and inferring dynamic features. In terms of CLT, this should correspond to a decrease in ECL, thereby leaving resources available that can be devoted to GCL. With respect to learning outcomes, these benefits of dynamic visualizations should be best reflected in performance on transfer tasks. With regard to learner characteristics, one may assume that the benefits of dynamic visualizations are especially pronounced for learners with weaker spatial abilities, and only to a lesser extent for learners with stronger spatial abilities. These abovementioned assumptions will be examined in Study 1.

Furthermore, to better understand learning from text, static and dynamic visualizations, and the functions the visualizations may serve for, process data might be particularly helpful, because they might give deeper insights in the associated cognitive processes when learning with this kind of multimedia instruction. At this, assumptions like the one of a more elaborate processing when visualizations are added to text can be examined more directly. This topic of assessing and analyzing cognitive processes will be considered in more detail in Study 1 of the current thesis.

However, even though dynamic visualizations are supposed to be best suited to convey the content of the domain at hand, a potential drawback of dynamic visualizations, namely their visual complexity, might still be apparent in Study 1. To further optimize learning with dynamic visualizations, it might be necessary to counteract the potential drawbacks arising from visual complexity. That is on the one hand the inter-representational split-attention effect, which might be especially problematic for visualizations that possess a high degree of visual complexity, and on the other hand the visual complexity within dynamic visualizations. This issue will be taken up again after Study 1.

3 Study 1: Can Differences in Learning Strategies Explain the Benefits of Learning From Static and Dynamic Visualizations?[7]

The first goal of Study 1 was to establish the multimedia effect for the current set of materials by determining that adding visualizations to text, compared to only text, helps to gain a deeper understanding of the domain. If the multimedia effect would not be given and text was sufficient for an adequate understanding of this kind of learning material, there would hardly be a reason to expect differences in learning with dynamic and static visualizations accompanied by text. A second goal was to investigate whether dynamic visualizations would be more apt for learners to achieve a deeper understanding than static visualizations, and if spatial abilities would moderate learning with dynamic and static visualizations. A third goal was to get further insights in the cognitive processes when learning with the used instructional material, that is either text, or text and dynamic visualizations, or text and static visualizations. To achieve this goal, think-aloud protocols (cf. Ericsson & Simon, 1993), as well as behavioral data during learning were assessed.

When applying the think-aloud procedure, participants are asked to think aloud while *concurrently* dealing with learning content (Ericsson & Simon, 1993). According to Ericsson and Simon, the think-aloud procedure is not supposed to interfere with the learning process itself as long as decisive guidelines are followed, such as not prompting the learners to verbalize certain aspects of the content. When assessing cognitive processes via verbal protocols, these verbal data have to be coded in order to classify the cognitive processes that took place. A systematic account of the different cognitive processes that can take place during learning with different multimedia instruction has been provided by Weinstein and Mayer (1986) in their seminal work on learning strategies, which also underlies the active processing assumption of the CTML. According to Weinstein and Mayer (1986, p. 315), learning strategies "can be defined as behaviors and thoughts that a learner engages in during learning and that are intended to influence the learner's encoding process". The quality of this encoding process is then supposed to affect learning outcomes. Weinstein and Mayer (1986), among other things, distinguish between rehearsal and elaboration strategies, as well as meta-cognitive strategies, such as comprehension monitoring strategies. Rehearsal strategies can be regarded as strategies where learners mainly recapitulate what was explicitly displayed in the multimedia instructions and that contribute to the

[7] This chapter is based on: Kühl, T., Scheiter, K., Gerjets, P., & Gemballa, S. (2011). Can differences in learning strategies explain the benefits of learning from static and dynamic visualizations? *Computers & Education, 56*, 176-187.

memorization of facts, but not to a deeper understanding. Elaboration strategies on the other hand are supposed to be involved in the construction and integration of new information, and should contribute to a deeper understanding of the content. Elaboration strategies can comprise processing activities such as connecting new content to already existing knowledge, as well as reasoning and inferring content that is not explicitly conveyed through the multimedia instruction. Comprehension monitoring strategies comprise activities that refer to a checking of one's own understanding of the multimedia instruction, which can be either positive or negative. This learning strategy classification laid the foundation for the formulation of the CTML's active processing assumption (see also Kombartzky, Ploetzner, Schlag, & Metz, 2010), and provides useful categories for distinguishing among different cognitive processes during learning. In line with this reasoning, the classification of cognitive processes for verbal data as learning strategies was successfully implemented in a study by Lewalter (2003). A similar coding scheme was used by Butcher (2006), with the categories paraphrases, which is essentially the same as rehearsal strategies, elaborations, self-explanation inferences and monitoring statements. Additionally, Butcher used the category errors (i.e., erroneous statements), which was also incorporated in the coding scheme of Study 1, as it may yield insight into misconceptions that arise during learning. More details about the coding scheme are provided in the method section of Study 1.

By assessing think-aloud protocols, often articulated assumptions – as derived from the functional view on learning with multimedia (cf. Chapter 2.1.2) – about learning with text and visualizations as opposed to text might be investigated more directly. Moreover, with regard to the cognitive processes arising when learning with dynamic and static visualizations, there is hardly any research investigating the articulated claims more directly by means of think-aloud protocols (for instance such as an illusion of understanding in learning with dynamic visualizations). More precisely, to the author's knowledge there is only one study (Lewalter, 2003) in which think-aloud protocols were used to assess the cognitive processes associated with learning with these types of visualizations in a multimedia context. The results of this study revealed that learners in the static compared to the dynamic visualization condition more often reproduced what had been described in the learning environment (rehearsal), and that this approach generally contributed to a better performance in (verbal) factual knowledge tasks, but not in transfer tasks. Also, learners in the static visualization condition more often made statements regarding the planning and regulation of further steps in learning, which was associated with a higher performance in transfer tasks. Lastly, learners in the dynamic visualization condition more often stated that they had understood the content (positive monitoring), even though this was not the case.

It should be noted though that participants in Lewalter's study were only asked to think aloud while watching visualizations, but not while reading the text. However, in Study 1 of this thesis, a topic of interest concerned the cognitive processes of participants while reading text and the interplay between text and visualizations as opposed to text alone, which already points to one crucial difference with regard to the research questions between the study of Lewalter (2003) and Study 1. Another, rather methodological difference between these studies was that in Study 1 of this thesis, a goal was to investigate these processes under more ecologically valid conditions: For instance, whereas in the study by Lewalter, text and visualizations were each presented on separate pages, never next to each other, in the current Study 1 text and visualizations were presented simultaneously. Moreover, in Lewalter's study learners had no opportunity to navigate backwards in order to re-examine the multimedia material. In contrast, in the Study 1 learners could navigate back and forth through the learning environment, and decide by themselves when and how often to watch the visualizations.

3.1 Hypotheses and Research Questions

First, the multimedia effect was expected to apply. With respect to learning outcomes, learners provided with text with dynamic or static visualizations should outperform learners provided with only text. This superiority should be more pronounced for pictorial tasks as well as transfer tasks as opposed to verbal factual knowledge tasks. With regard to processing demands, adding visualizations to text should offload working memory as compared to receiving only text, and thereby decrease ECL. Moreover, it should also lead to a more thorough processing of the content, and, hence, in an increase in GCL. Concerning the assessed cognitive processes as coded by learning strategies, a more thorough processing of the content might be reflected by more elaborations. Also, in line with the functional view that was outlined in Chapter 2.1.2, it was assumed that learning with text alone as opposed to learning with text and visualizations would lead to more uncertainty about the domain. With regard to learning strategies, this uncertainty might be reflected by more erroneous statements (cf. Butcher, 2006), as well as more negative monitoring statements, and on the other hand to fewer positive monitoring statements.

Second, for the domain of the study at hand, dynamic visualizations were expected to be superior to static visualizations with respect to learning outcomes, and specifically transfer tasks. Correspondingly, with respect to processing demands, dynamic visualizations were expected to decrease ECL. As was explicated in Chapter 2.2.2, inferring spatial and temporal changes with static visualizations may be an error-prone process (Schnotz & Lowe, 2008). Concerning learning strategies, this might be reflected in comparatively more erroneous statements as well as more negative monitoring statements regarding the content when learning with static visualizations. Moreover, it was assumed that learners with dynamic visualizations would produce comparatively more positive monitoring statements, fewer rehearsal statements, and fewer statements about the planning and regulation of further steps in learning (see also Lewalter, 2003).

Third, according to the ability-as compensator hypothesis, spatial abilities were expected to moderate the effectiveness of learning with these two different types of visualizations: The superiority of dynamic over static visualizations should be even more pronounced for learners with weaker spatial abilities than for learners with stronger spatial abilities. Alternatively, as mentioned in Chapter 2.2.4.2, learners in the static visualization condition might attempt to compensate for the drawbacks of static visualizations, and might therefore decide to watch the visualizations more often to understand the dynamic content. Again, this may be especially the case for learners with weaker spatial abilities.

3.2 Method

3.2.1 Participants and Design

Seventy-five students with various educational backgrounds from the University of Tuebingen, Germany, participated in the study in return for either course credit or payment. Due to technical problems, the incomplete data from three participants had to be excluded. The remaining 72 students were 45 female and 27 male participants (average age M = 24.32 years, SD = 3.10). The design comprised three conditions: a text-only condition (TOC), a condition that combined text with dynamic visualizations (DVC), and a condition that combined text with static visualizations (SVC), at which 24 participants served in each condition.

3.2.2 Instructional Materials

The computerized instructional material dealt with the physical principles underlying undulatory fish locomotion. The learning environment consisted of an introduction, where the concept of a parallelogram of forces was explained, as well as of a learning phase. The introduction contained 250 words and four static visualizations that illustrated the different steps in constructing a parallelogram of forces, and was the same for all three conditions. Subsequently, the learning phase began, which was subject to experimental manipulation. The expository text (1,513 words) was adequately comprehendible that the average learner could answer all questions after reading. The text of the learning phase was distributed across thirteen pages. Learners could navigate through the learning environment by clicking a “Next”-button or a “Back”-button, respectively. In all conditions, the same written text was presented on the left half of the screen. Written text as well as self-pacing was implemented to ensure that think-aloud protocols could be assessed properly. For participants of the text-only condition, the right half of the screen was blank.

For the DVC, dynamic visualizations were presented on the right half of the screen, and the first frame of a dynamic visualization was presented on the screen until the visualization was started by clicking a “Play”-button (see Figure 3.1). To ensure that learners were playing the whole dynamic visualization, the “Next”-button of a page appeared only after a dynamic visualization had ended. When the dynamic visualization ended, the last frame remained visible on the screen. Learners could then either replay the visualizations (and/or read the text, respectively) or move on to the next page. Except for pressing the “Back”-, “Next”-, and “Play”-button, no further interactivity was implemented in the learning environment. The DVC consisted of one photorealistic video and eleven computer-generated animations (cf. Figure 3.1). The video was displayed at the beginning of the learning phase, and showed various fish applying different locomotion patterns that have developed in the course of evolution. The computer-generated animations, which contained both, rather photorealistic as well as more schematized illustrations, depicted the underlying physical principles of an undulatory (i.e., wave-like) swimming fish. At this, they displayed translations, transitions, and transformations (cf. Lowe, 2003, 2004). More precisely, these computer-generated animations conveyed information about the interplay of the trajectory and velocity of the body parts, the corresponding displacement of water, as well as the relation between these variables, and the size of the associated resulting forces and their direction. These forces were represented as arrows, and continuously varied in length and spatial orientation depending on the force’s strength and direction, and can be regarded as an example for transformations. An example for translations would be the movement of the caudal fin, and for transitions, an example would be the displaced water that disappears from the visualization. Moreover, the animations depicted how the different forces that act upon different segments of the fish body add up or cancel each other out, respectively. Sometimes, in crucial states, the animation stopped and verbal labels for these states appeared.

In contrast to the DVC, the SVC consisted of a sequence of extracted frames from the corresponding dynamic visualizations. In case labels appeared in the dynamic visualizations, the frames of the SVC also comprised the same labels to keep the static visualizations as similar as possible to the dynamic visualizations regarding the available information. These frames were presented sequentially one after the other once a learner had clicked on the "Play"-button, and contained between nine to sixteen frames depending on the corresponding dynamic visualization. The condition with static visualizations was identical to the DVC in terms of navigation, interactivity, resolution, and size of the visualizations. Also, the duration of a sequence of static visualizations was identical to the duration of the corresponding dynamic visualizations.

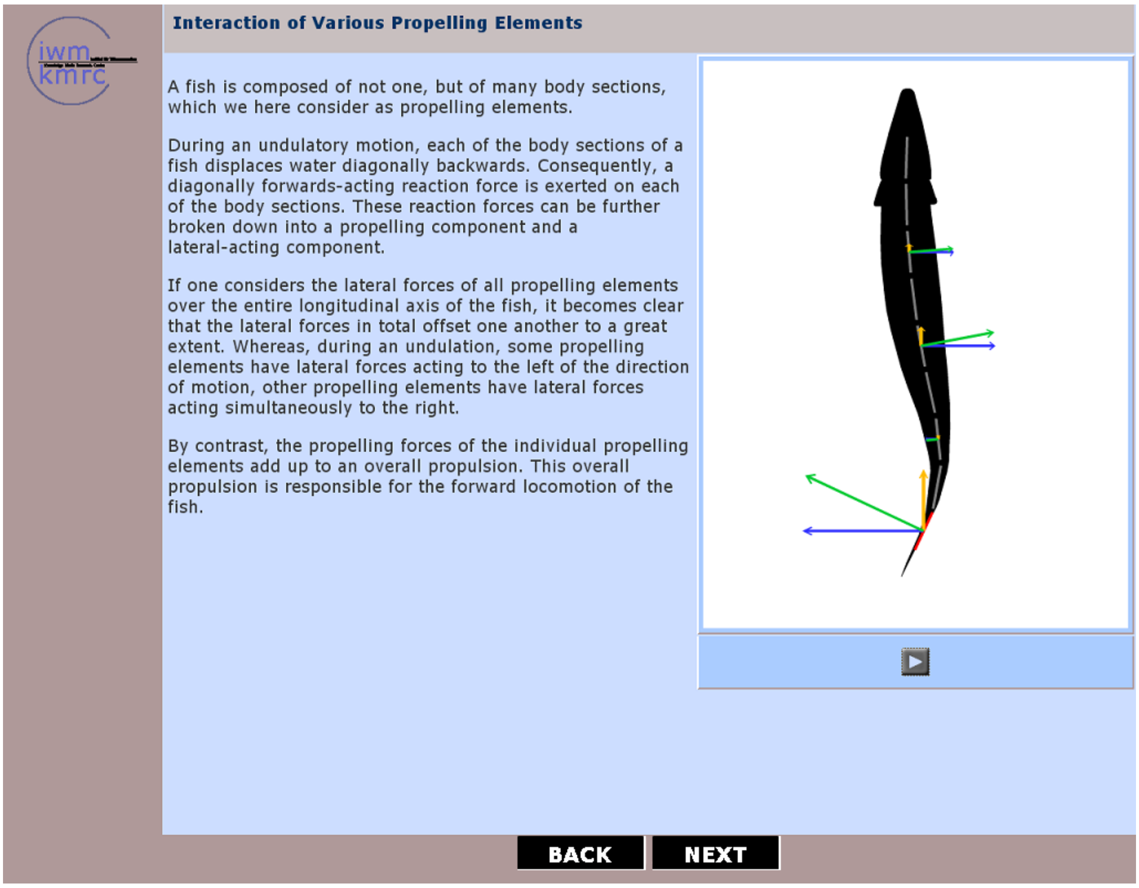


Figure 3.1. Snapshots of the learning environment for the visualization conditions (original text in German).

3.2.3 Measures

A questionnaire with respect to the attitudes towards biology and physics, as well as a prerequisite knowledge test served as control variables, and a spatial ability test served as a moderator variable (and also as a control variable). The dependent measures consisted of variables addressing the use of the learning environment (i.e., learning time and frequency of playing the visualizations), statements derived from the think-aloud protocols, cognitive load items, and several learning outcome measures.

Attitudes towards biology and physics. The questionnaire was an adapted version of an attitude scale towards biology by Russell and Hollander (1975). For seven of the fourteen items of that attitude scale, the word “biology” was substituted by “physics”, with the aim of not only being able to measure attitude towards biology, but also to measure attitudes towards physics. The fourteen items of that questionnaire had to be rated on a 5-point Likert scale, ranging from 1 (“I strongly disagree”) to 5 (“I strongly agree”). The wording of the fourteen questions is provided in Appendix A (in their original German version). For further analysis, the negative formulated items were recoded, so that the higher a score is on the scale, the more positive a participant’s attitude is towards biology or physics, respectively.

Prerequisite knowledge. The prerequisite knowledge test consisted of 5 multiple-choice questions asking for the second and third Newton axioms, the physical definition of forces, the characteristics of a harmonic oscillation, and knowledge about velocity and acceleration (see sample item below). A person’s knowledge about these basic definitions and principles was considered a beneficial prerequisite for more easily achieving an understanding of the topics explained in the current study. Hence, it was not the aim of testing for a deeper understanding of physical concepts or principles that were unlikely to be present in the used student sample. For each correct answer to a question, learners were assigned one point, and for each wrong answer, one point was subtracted. Within a question, however, learners could receive a minimum of zero points, resulting in a maximum of nine points (3 items with 1 correct answer, 2 items with 3 correct answers).

Example of a question from the prerequisite knowledge test

According to Newton’s second law of motion, a force *F* is calculated from

a) the product of mass and time

b) the product of mass and acceleration

c) the product of time and impulse

d) the product of impulse and acceleration

Spatial ability. To control for individual differences in spatial abilities, and to examine their potential moderating role, the mental rotation test (MRT) was administered (Vandenberg & Kuse, 1978). As especially learning with the static visualizations used in this study required the ability to mentally rotate and manipulate spatial objects (e.g., to imagine the movement of the caudal fin), the MRT – which loads on the factor VZ of spatial abilities – was applied, since the MRT was assumed to fit well to these requirements. The MRT consists of 20 items, whereby each item comprises a complex three-dimensional block figure and four alternative figures as multiple-choice answer options. For each item, the participant has to choose, which two of the four alternative figures are identical to the target when (mentally) rotated. There is a time limit of six minutes for working on the MRT. For each correctly identified figure one point was given, and for each wrong identified figure one point was subtracted, resulting in a maximum of 40 points and a minimum of -40 points.

Use of the learning environment. Because learners could decide how long they wanted to learn, learning time served as dependent variable. Furthermore, as learners in the two illustrated conditions could decide on how often to play the visualizations, the frequency of playing the visualizations was registered as another dependent variable. Due to recording problems, the data analyses of the use of the learning environment could only be conducted with the data of 71 instead of 72 participants.

Think-aloud protocols. The categories for coding the protocols were built on an adapted version of Butcher's (2006) and Lewalter's (2003) coding schemes, and were refined by analyzing sub samples of the protocols. The segmentation of the protocols was at a small grain size, in which sentences, subordinate clauses, or utterances preceded and followed by a pause were considered as separate segments (cf. van Gog, Paas, van Merriënboer, & Witte, 2005). Fifteen protocols were coded independently by two raters according to the refined coding scheme, with an inter-rater reliability of .73 (Cohen's kappa). As inter-rater reliability was considerably good (van Someren, Barnard, & Sandberg, 1994), one rater who was blind with respect to the research questions coded the remaining protocols, and only the coding of this rater was used for further data analyses. To code the protocols, the software tool MEPA 4.10 (Erkens, 2005) was used. For each participant, the number of codes for each category was counted. Due to recording problems, the data analyses of the think-aloud protocols could only be conducted with the data of 71 instead of 72 participants.

The main categories were rehearsal, elaboration, monitoring, erroneous statements, and planning for further learning (cf. Butcher, 2006; Lewalter, 2003). The category *rehearsal* referred to statements, which solely reproduced what had been explicitly described in the learning environment. The category *elaboration* was divided in two subcategories, namely activation of knowledge and generative inferences. The subcategory *activation of knowledge* referred to statements, which showed strategies of linking the content to prior knowledge, or to what had already been learned from the instructional environment, whereas the subcategory *generative inferences* comprised statements, which referred to new deeper insights about the content and went beyond what was covered by the instructional material studied at the time the statement was made. The category *monitoring* referred to statements, which showed an evaluation of the actual learning process, which was either judged as *positive* or as *negative*. The category *erroneous statements* consisted of wrong reproductions or faulty elaborations. The category *planning for further learning* referred to statements aiming at the planning for further steps in learning[8]. Table 3.1 provides examples of the abovementioned categories.

Table 3.1

Coding Scheme of the Categories and Their Respective Examples

Categories	Examples
Rehearsal	"every body section is considered as a propelling element"
Activation of knowledge	"when I swim, I also try to push myself off the water", or "as was mentioned before, the reaction force is dependent of the acceleration and height of the moving fish"
Generative inferences	"if the lateral forces would not cancel each other out, the fish would swim diagonally and not straight on"
Positive monitoring	"ok, this is clear"
Negative monitoring	"I do not understand!"
Erroneous statements	"ok, at the zero baseline, the velocity of a propelling element is lowest"
Planning for further learning	"now I just will watch the animation, maybe this will help", or "I have to read this again"

Cognitive load measures. To be able to distinguish between ECL and GCL, two items were used to measure cognitive load after the learning phase. The item *perceived difficulty of the tasks* ("How difficult was it for you to understand the contents?") was supposed to measure ECL, whereas the item *mental effort* ("How much effort did you invest in order to understand the content?") was supposed to measure GCL (cf. Gerjets, Scheiter, Opfermann, Hesse, & Eysink, 2009). Each item had to be rated by the participant on a scale ranging from 1 to 21.

Knowledge tests. Learning outcomes were measured by means of verbal factual knowledge tasks (thirteen multiple-choice questions), three pictorial recall tasks, and eleven transfer tasks (see sample items of each test below). A maximum of 25 points could be achieved for the verbal factual knowledge test, a maximum of 6 points could be achieved for the pictorial recall test, and a maximum of 29 points could be achieved for the transfer test. All correct

[8] Statements about the design of the learning environment, as well as statements that could not be assigned to any category (e.g., program-related questions or isolated expressions such as "ok"), were subsumed to a rest category, which, however, will be neglected in the following.

answers to the verbal factual knowledge tasks had been explicitly mentioned in the text of the instructional material. The pictorial recall tasks were posed in pictorial format, and required the participant to work with pictures. The correct answer had been described in the text, and could additionally be seen in the two visualization conditions. The eleven transfer tasks were posed in verbal as well as in pictorial form. To solve the transfer tasks, learners had to apply what they had learned to new situations and problems. The transfer questions can be considered near transfer tasks in that the questions always referred to the situation of objects moving in water (e.g., fish, boats) under varying conditions (e.g., moving backwards). Hence, the context in which the questions were embedded was identical to that of the learning materials, but they required modifications of what had been learned to accommodate the task requirements (cf. Barnett & Ceci, 2002).

The pictorial recall tasks as well as the transfer tasks were scored by two independent raters. For pictorial recall, raters agreed on 90%, and for transfer tasks, raters agreed on 95% of the given answers. Cases of disagreement were resolved by reaching a consensus.

Example of a question from the verbal factual knowledge test

Which of the following is/are true?

a) The reaction force acts in the opposite direction to the motion of the propelling element.

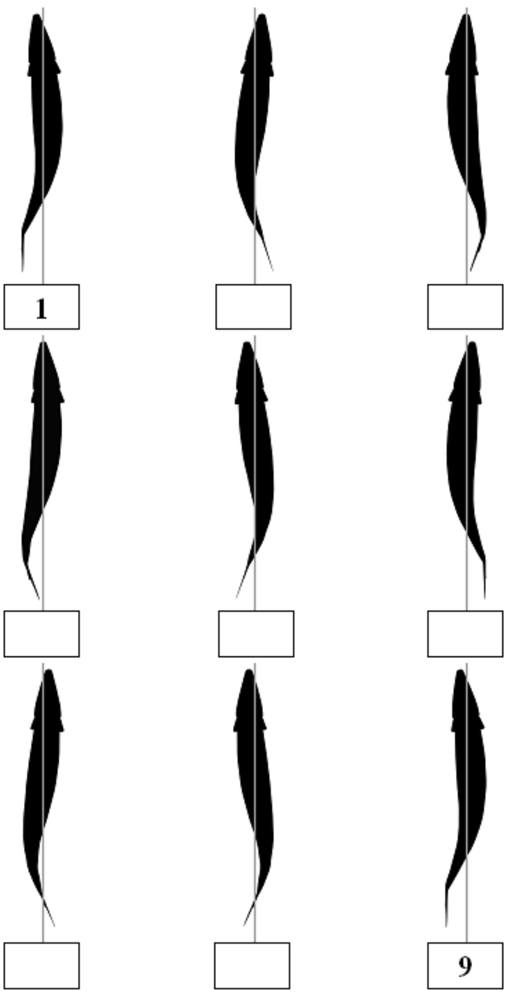
b) The reaction force acts in the direction of motion of the propelling element.

c) The reaction force forms a right angle to the propelling element.

d) The reaction force forms a right angle to the swimming direction.

Example of a question from the pictorial recall test

Arrange the different states of the fish movement in the correct order by writing the corresponding numbers in the respective boxes.



Example of a question from the transfer test

Some undulating species of fish move their head back and forth in order to swim forwards. Why is this? Write down any feasible reasons you can think of!

3.2.4 Procedure

Each participant was tested individually in a session lasting between 90 and 120 minutes. First, the participants received a written overview of the procedure, which was followed by the questionnaire concerning the attitude towards biology and physics as well as by the prerequisite knowledge test. Thereafter, they were required to work on the mental rotation test with a time limit of six minutes. Then, before they started studying the instructional materials, each participant was asked to practice thinking aloud while reading the introduction on the parallelogram of forces. The think-aloud protocols were recorded with Camtasia 3.0. Subsequently, the learning phase began. The participants could use the instructional material without any time constraints, and were asked to think aloud the whole time. When they stopped talking for more than 15 seconds they were prompted to think aloud. Having finished learning, they were provided with the cognitive load items to measure processing demands during the learning phase. Afterwards, learning outcomes were assessed by the knowledge tests.

3.3 Results

In a first step, the questionnaire concerning attitudes towards biology and physics was analyzed by means of a factor analysis, to validate the assumed independence of the two factors. Then students' attitudes towards biology and physics, their prerequisite knowledge, as well as their spatial ability scores were analyzed by means of ANOVAs to test if the experimental conditions could be regarded as equal with respect to these influencing variables. In a second step, specific contrasts were used to test the specified hypotheses. Therefore, firstly the text-only condition was compared to the two visualizations conditions, and secondly, the hypotheses concerning dynamic and static visualizations for the dependent variables learning outcomes, cognitive load, and learning strategies were tested. Furthermore, it was tested if spatial abilities moderated learning with dynamic and static visualizations with regard to learning outcomes, as well as the frequency of the usage of visualizations in the two visualization conditions. For this purpose, the scores of the spatial ability test for participants of the two visualization conditions were z-standardized and used as a continuous factor in the respective ANCOVAs. In case an interaction between spatial ability and the two visualization conditions occurred, correlations were computed to determine the direction of this interaction. Partial eta-squared ($\eta^2 p$) is reported as measures of effect size.

3.3.1 Comparability of Experimental Conditions with Respect to Attitude Towards Biology and Physics, Prerequisite Knowledge, Spatial Abilities, and Learning Time

Because two different scales were assumed for the attitude questionnaire, namely a biology scale, and a physics scale, two factors were extracted by principle component analysis and rotated by varimax rotation. Note that negatively formulated items were recoded. Loading of items on factors are depicted in Appendix A. Items are grouped by factors and by size of loading to facilitate interpretation. As can be seen in Appendix A, the items loaded on the one hand well on their assumed factors with all loadings > .60 (cf. Bortz, 2005), and, moreover, poorly on the other factor (all loadings < .30), indicating that the biology scale and the physics scale are two independent constructs.

In a next step, the internal consistency was determined for each of the two scales. For the biology scale (item 1-7) Cronbach's alpha (α) was α = .90, and the physics scale (item 8-14) revealed α =.94. Due to the high internal consistency of each scale, the items of each scale were subsumed to one biology score and one physics score, respectively. Means and standard deviations for attitudes towards biology as well as physics, prerequisite knowledge, and spatial ability are reported in Table 3.2. There were no statistically significant differences between the three instructional conditions concerning either attitude towards biology, attitude towards physics, prerequisite knowledge (all *Fs* < 1, *ns*), or spatial abilities ($F(2, 69) = 1.77$, $MSE = 72.60$, $p = .18$, $\eta^2 p = .05$), so that the conditions could be regarded as equal with respect to these variables.
Data concerning learning time are depicted in Table 3.2. A one-factorial ANOVA for learning time revealed an overall effect for the three instructional conditions ($F(2, 68) = 5.13$, $MSE = 37.86$, $p < .01$, $\eta^2 p = .13$). Bonferroni post-hoc tests revealed significant longer learning times for the SVC compared to the TOC ($p < .01$), but no differences between the DVC compared to the TOC ($p = .12$) nor between the two illustrated conditions ($p = .84$). Because of these differences, learning time was used as a covariate in the analyses of learning outcomes.

Table 3.2

Means (and SD) for Control Variables, Use of Learning Environment, Learning Outcomes, and Cognitive Load

	Text (n = 24)	Static (n = 24)	Dynamic (n = 24)
Attitude towards biology-scale (7-35)	29.75 (5.85)	28.92 (5.26)	29.25 (5.67)
Attitude towards physics-scale (7-35)	18.29 (7.37)	19.88 (8.50)	19.04 (8.35)
Prerequisite knowledge (%)	56.94 (22.54)	54.63 (24.94)	57.87 (23.39)
Spatial abilities	18.00 (6.19)	21.88 (10.12)	17.75 (8.78)
Use of learning environment [a]			
Learning time (in minutes)	20.98 (7.48)	26.63 (5.65) [a]	24.67 (5.05)
Frequency of using visualizations	-	15.70 (3.90) [a]	13.95 (2.18)
Learning outcomes [b]			
Factual knowledge (%)	57.52 (3.64)	61.38 (3.67)	59.16 (3.49)
Pictorial recall (%)	39.06 (5.78)	59.35 (5.83)	60.31 (5.55)
Transfer (%)	38.98 (3.11)	50.39 (3.13)	52.18 (2.98)
Cognitive load (1-21)			
ECL	14.38 (4.23)	10.75 (5.11)	10.71 (4.15)
GCL	15.92 (3.61)	15.45 (3.65)	15.58 (3.15)

[a] Only the data of 23 participants were available due to recording problems.
[b] Note: Learning outcomes are adjusted by taking learning time into account; values in parentheses refer to standard errors for this dependent measure. This leads to the exclusion of one subject, whose learning time was not available due to recording problems.

3.3.2 Effects of Adding Visualizations to Text

The three experimental conditions were compared with regard to their instructional effectiveness by one-factorial ANCOVAs with the dependent variables verbal factual knowledge, pictorial recall, and transfer knowledge, and learning time acting as a covariate (cf. Table 3.2). There was no significant influence of learning time on verbal factual knowledge or pictorial recall (both *Fs* < 1, *ns*), but a marginal significant effect on transfer ($F(1, 67) = 3.04$, $MSE = 646.18$, $p = .09$, $\eta^2 p = .04$). However, learning time did not correlate significantly with transfer ($N = 71$; $r = -.09$; $p = .47$)[9].

Planned contrasts for learning outcomes concerning differences between the text-only and the two visualization conditions with learning time as a covariate, revealed no differences for factual knowledge ($F < 1$, *ns)*, but for pictorial recall ($F(1, 67) = 8.19$, $MSE = 737.47$, $p < .01$, $\eta^2 p$ =

[9] It should be noted that the same pattern of results for the different learning outcome measures was observable when learning time was not considered as a covariate.

.11) as well as for transfer ($F(1, 67) = 9.97$, $MSE = 212.71$, $p < .01$, $\eta^2 p = .13$), indicating that, in line with the first hypothesis, for the latter two tests learners provided with visualizations outperformed learners provided with only text.

For cognitive load, planned contrasts between the two visualization conditions and the text-only condition revealed a significant effect for ECL ($F(1, 69) = 10.41$, $MSE = 20.42$, $p < .01$, $\eta^2 p = .13$), with learners in the text-only condition perceiving the content as more difficult than learners in the visualization condition, while there were no differences for GCL ($F < 1$, *ns*).

Means and standard deviations for the categories of the think-aloud protocols are depicted in Table 3.3[10]. Due to the fact that the think-aloud data violated the assumption of normal distribution, non-parametric Mann-Whitney's U-statistics were applied. Planned contrasts between the two visualization conditions and the text-only condition for elaborations, namely for the subcategories generative inferences and activation of knowledge, revealed a significant effect for generative inferences ($U = 365$, $p = .02$), but not for statements concerning the activation of knowledge ($U = 533.5$, $p = .71$). As predicted, learners in the two visualization conditions generated more inferences ($Mdn_{VIS} = 6.00$) than learners in the text-only condition ($Mdn_{TOC} = 3.50$), whereas however, this was not the case for activation of knowledge ($Mdn_{VIS} = 6.00$ and $Mdn_{TOC} = 5.00$, respectively). Comparing the two visualization conditions and the TOC revealed no effect for positive monitoring ($U = 524$, $p = .63$), but a significant effect for negative monitoring ($U = 364$, $p = .02$). As predicted, learners in the two visualization conditions less often stated that they did not understand the content ($Mdn_{VIS} = 4.00$) compared to learners in the TOC ($Mdn_{TOC} = 6.00$), whereas on the other hand no differences for positive monitoring emerged ($Mdn_{VIS} = 5.00$ and $Mdn_{TOC} = 6.50$, respectively). Also, no differences were observable between the visualization conditions and the TOC with respect to erroneous statements ($U = 536.5$, $p = .73$; both $Mdn = 3.00$).

To sum up, learners in the conditions receiving text and visualizations as opposed to the text-only condition performed better on pictorial tasks and transfer tasks, perceived the learning task as less difficult, generated more inferences and less often stated that they did not understand the content.

[10] Kruskal-Wallis' tests revealed that the overall number of codes as well as the number of codes per minute learning time were comparable across conditions (overall number of codes: $H(2) = 2.67$, $p = .27$; number of codes per minute learning time: $H(2) = 0.15$, $p = .93$).

Table 3.3

Means (and SD) for Processing Activities

	Text (n = 24)	Static (n = 23) [a]	Dynamic (n = 24)
Processing activities			
Overall number of codes	95.71 (46.97)	115.30 (53.38)	112.88 (37.50)
Number of codes per learning time (in minute)	4.67 (1.97)	4.51 (2.25)	4.58 (1.25)
Rehearsal	40.63 (34.20)	43.26 (36.60)	33.21 (23.69)
Activation of knowledge	5.75 (4.96)	7.22 (6.45)	5.00 (3.38)
Generative inferences	4.54 (5.18)	7.52 (6.19)	8.38 (6.67)
Positive monitoring	8.21 (7.35)	4.52 (4.21)	9.42 (7.92)
Negative monitoring	6.83 (3.64)	4.48 (4.02)	5.17 (5.11)
Erroneous statements	3.42 (3.65)	3.48 (2.37)	2.46 (3.72)
Planning for further learning	1.83 (2.24)	3.00 (3.83)	4.25 (5.29)

[a] Only the data of 23 participants were available due to recording problems.

3.3.3 Effects of Dynamic versus Static Visualizations

With regard to learning outcomes, planned contrasts between the dynamic and static visualization condition, with learning time as a covariate, revealed no differences for factual knowledge, for pictorial recall, or for transfer (all $Fs < 1$, *ns*).

Concerning cognitive load, differences between the two visualization conditions emerged neither for ECL nor for GCL (both $Fs < 1$, *ns*).

With respect to learning strategies, planned contrasts between the DVC and the SVC revealed no significant differences between the conditions concerning activation of knowledge ($U = 223.5$, $p = .26$; $Mdn_{DVC} = 5.00$ and $Mdn_{SVC} = 6.00$, respectively), generative inferences ($U = 254$, $p = .64$; $Mdn_{DVC} = 6.00$ and $Mdn_{SVC} = 5.00$, respectively), negative monitoring ($U = 262$, $p = .76$; both $Mdn = 4.00$), rehearsal ($U = 246.5$, $p = .53$; $Mdn_{DVC} = 28.50$ and $Mdn_{SVC} = 35.00$, respectively) and planning for further learning ($U = 247.5$, $p = .54$; both $Mdn = 2.00$). However, learners in the DVC made more positive monitoring statements than learners in the SVC ($U = 156$, $p = .01$; $Mdn_{DVC} = 8.00$ and $Mdn_{SVC} = 3.00$, respectively) and less erroneous statements ($U = 168$, $p = .02$; $Mdn_{DVC} = 1.00$ and $Mdn_{SVC} = 3.00$, respectively).

In sum, there were no differences between the DVC and the SVC with respect to learning outcomes or subjectively rated cognitive load. Concerning learning strategies, more positive monitoring statements as well as less erroneous statements were conducted in the DVC compared to the SVC.

3.3.4 The Moderating Role of Spatial Abilities

To test the moderating role of spatial abilities for learning with the two different visualization formats, one-factorial ANCOVAs with the different learning outcome measures as dependent variables and spatial abilities as a continuous factor was conducted. Concerning type of visualizations, the one-factorial ANCOVAs revealed neither differences for factual knowledge ($F < 1$, *ns*), nor for pictorial recall ($F < 1$, *ns*), nor for transfer ($F(1, 44) = 2.18$, $MSE = 224.76$, $p = .15$, $\eta^2 p = .05$). Also with respect to the moderation of spatial abilities with type of visualizations, one-factorial ANCOVAs revealed neither an interaction for factual knowledge ($F < 1$, *ns*), nor for pictorial recall ($F(1, 44) = 2.21$, $MSE = 690.45$, $p = .14$, $\eta^2 p = .05$), nor for transfer ($F < 1$, *ns*). However, the one-factorial ANCOVAs revealed a significant main effect of spatial abilities for pictorial recall ($F(1, 44) = 8.21$, $MSE = 690.45$, $p < .01$, $\eta^2 p = .16$) and transfer ($F(1, 44) = 11.16$, $MSE = 224.76$, $p < .01$, $\eta^2 p = .20$), and a marginally significant effect for factual knowledge ($F(1, 44) = 3.07$, $MSE = 284.71$, $p = .09$, $\eta^2 p = .07$). The effects indicated that higher spatial abilities were associated with better performance in pictorial recall (N = 48; $r = .35$; $p = .01$) as well as transfer (N = 48; $r = .42$; $p < .01$), and marginally associated with better performance in factual knowledge (N = 48; $r = .26$; $p = .08$).

To test whether differences in spatial abilities were associated with differences in the frequency of using different visualization formats, a one-factorial ANCOVA with the frequency of using visualizations as dependent variable and spatial abilities as a continuous factor was conducted. No main effect could be observed for spatial abilities ($F < 1$), but a significant effect for instructional condition could be observed ($F(1, 43) = 4.37$, $MSE = 9.16$, $p = .04$, $\eta^2 p = .09$), with learners in the static condition using the visualizations more often. Moreover, a significant interaction between spatial abilities and instructional condition ($F(1, 43) = 4.33$, $MSE = 9.16$, $p = .04$, $\eta^2 p = .09$) was found, indicating that spatial abilities moderated the frequency of the usage of visualizations in learning with dynamic and static visualizations. For the DVC, no significant correlations between spatial abilities and usage of visualizations were observable (N = 24; $r = .26$; $p = .23$), while for the SVC, there was a marginally significant relationship between spatial abilities and usage of visualization (N = 23; $r = -.36$; $p = .10$), meaning that the lower the spatial abilities were for learners in the SVC the more often they used the visualizations.

To test the moderating role of spatial abilities on cognitive load with the two different visualization formats, one-factorial ANCOVAs with the two different cognitive load measures as dependent variables and spatial abilities as a continuous factor was conducted. Concerning type of visualizations, the one-factorial ANCOVAs revealed no differences for ECL, or for GCL (both *Fs* < 1, *ns*). Also with respect to the moderation of spatial abilities with type of visualizations, one-factorial ANCOVAs revealed no interaction for ECL ($F < 1$, *ns*), or for GCL ($F(1, 44) = 1.06$, $MSE = 11.82$, $p = .31$, $\eta^2 p = .02$). Furthermore, no main effect of spatial abilities could be observed for ECL ($F(1, 44) = 2.35$, $MSE = 21.54$, $p = .13$, $\eta^2 p = .05$), or for GCL ($F < 1$, *ns*).

In conclusion, spatial abilities did not moderate learning outcomes or cognitive load. However, they did moderate the frequency of playing the visualizations for the two visualization conditions in that the lower spatial abilities were for learners in the SVC, the more often they played the static visualizations, while there was no association between spatial abilities and the frequency of playing the visualizations for learners in the DVC.

3.3.5 Relationships between Learning Outcomes, Cognitive Load, and Learning Strategies

Correlations between the two cognitive load items and the different knowledge tasks (verbal factual knowledge, pictorial recall, transfer knowledge) were computed across conditions. There was always a negative relationship between ECL and each kind of knowledge task (verbal factual knowledge: $r = -.29$, $p < .05$; pictorial recall: $r = -.47$, $p < .01$; transfer tasks: $r = -.52$, $p < .01$), indicating that higher ECL was associated with lower learning outcomes. However, there were no relationships between GCL and pictorial recall ($r = .15$, $p = .20$), transfer ($r = -.04$, $p = .73$), or verbal factual knowledge ($r = .20$, $p = .10$).

The relationship between the categories of the think-aloud protocols and the different kind of learning outcomes are depicted in Table 3.4. Better factual knowledge was associated with more frequent verbalizations of rehearsal strategies, with more generative inferences, with less negative monitoring statements, and loosely associated with more frequent statements concerning the activation of knowledge. Pictorial recall was associated with more generative inferences and loosely associated with less negative monitoring statements. Similarly, transfer performance was related to fewer negative monitoring statements and slightly related to more generative inferences. There were no further relationships among the verbalizations and learning outcomes.

Table 3.4

Non-parametric Correlations (Kendall's τ) Among Categories of the Think-aloud Protocols and Knowledge Tasks

n = 71	Factual knowledge	Pictorial recall	Transfer knowledge
Rehearsal	$\tau = .20^*$	$\tau = .07$	$\tau = .05$
Activation of knowledge	$\tau = .15^{(*)}$	$\tau = .13$	$\tau = .04$
Generative inferences	$\tau = .17^*$	$\tau = .19^*$	$\tau = .14^{(*)}$
Positive monitoring	$\tau = .05$	$\tau = .09$	$\tau = .02$
Negative monitoring	$\tau = -.18^*$	$\tau = -.17^{(*)}$	$\tau = -.30^{**}$
Erroneous statements	$\tau = .04$	$\tau = -.06$	$\tau = -.11$
Planning for further learning	$\tau = -.04$	$\tau = -.04$	$\tau = -.05$

Note: (*) $p < .10$, * $p < .05$, ** $p < .01$

3.4 Summary and Discussion

In the current study, a text-only condition, a condition with text and static visualizations and a condition with text and dynamic visualizations were compared with respect to use of the learning environment, learning strategies, processing demands, and learning outcomes.

First, according to the multimedia effect, it was expected that the two visualization conditions would outperform the text-only condition for all learning outcome measures; this effect was assumed to be even more pronounced for pictorial recall and transfer tasks. Whereas for pictorial recall and transfer tasks the two visualization conditions outperformed the text-only condition, thereby confirming the first hypothesis, for verbal factual knowledge tasks however, no differences between these conditions were observed. A possible explanation for the latter might be found in the richness of the text. In the present study, the expository text was rich in detail to make it as "fair" as possible for the text-only condition, so that all relevant information that was depicted in the visualizations could in principle be inferred from the text (cf. informational equivalence, Larkin & Simon, 1987). Similarly, Mayer (2001) reported in his overview to have failed to find the multimedia effect for (verbal) factual knowledge in three of nine experiments. As he stated, the text for these three experiments was richer than in the other six experiments, where text and visualizations were somewhat complementary. Thus, maybe an enriched text, which is basically redundant to the visualizations, purges the multimedia effect for verbal factual knowledge. As expected, for accomplishing pictorial recall and transfer tasks, where a pictorial model of the content was assumed to be advantageous, the two visualization conditions outperformed the text-only condition. These results stress the importance of applying more differentiated learning outcome measures when investigating the instructional effectiveness of multimedia learning environments. With respect to learning time, it should be noted that even though learners provided with static visualizations dedicated more time to the instructional material compared to the text-only condition, their higher effectiveness was still existent when considering learning time as a covariate.

Regarding the abovementioned redundancy of text and visualizations, it should be noted that in some studies, the redundancy of text and visualizations might even be harmful for learning (cf. redundancy principle, Sweller, 2005b). For instance, in a recent study by Schmidt-Weigand and Scheiter (2011), the authors could not find a superiority of text and visualizations over text, if the text contained a high degree of spatial information; rather, there was a superiority of text and visualizations over text only when the text did contain a low degree of spatial information, so that text and visualizations had rather complementary roles. Moreover, learners receiving visualizations with text containing low degree of spatial information tended to outperform learners receiving text containing a high degree of spatial information. Hence, with respect to follow-up studies, from an instructional point of view, it seems recommendable to reduce the informational overlap between text and visualizations in a manner that advantages of text (e.g., conveying abstract knowledge) and advantages of visualizations (conveying visuo-spatial information) are emphasized. As a consequence, learning with different types of visualizations may also be more pronounced and may shine through, as learners may rely less on text, but have to rely more on the information depicted in the visualizations.

Concerning processing demands, it was expected that learners in the text-only condition would experience higher ECL than learners in the two visualization conditions. The score of the perceived difficulty item – which can be regarded as a candidate for measuring ECL, because it correlated negatively with all learning outcome measures – supported this hypothesis. In contrast, the self-reported mental effort item was not correlated with any of the performance measures. Hence, it is a less likely candidate for measuring GCL, which might also explain why there were no differences between conditions for this item. Problems in finding subjective measures that are suited to distinguish among different load types contribute to a growing body of research (cf. de

Jong, 2010). However, this does not mean that looking for such measures is not worthwhile, but rather that the search is not yet over.

With respect to learning strategies as measured by the verbal data of the think-aloud protocols, it was expected that fewer statements concerning the activation of knowledge, fewer generative inferences and fewer positive monitoring statements as well as more negative monitoring and more erroneous statements would occur in the text-only condition compared to the text and visualizations conditions. The verbal data gave partial support for these assumptions. On the one hand, learners in the text-only conditions did not produce fewer statements concerning the activation of knowledge, fewer positive monitoring, or more erroneous statements. However, it should be noted that these categories also had no substantial relation to learning outcomes. On the other hand, as predicted, learners in both visualization conditions stated less often that they did not understand the content, and, furthermore, conducted more generative inferences. Moreover, these categories had a substantial relation to learning outcomes, suggesting that they may be well suited in explaining the better learning outcomes of learners studying visualizations.

To sum up, these results indicate that adding visualizations to text can offload working memory, and encourage learners to engage in more valuable processing activities, which in turn results in a better understanding of the content.

Second, it was expected that dynamic visualizations would be more beneficial than static visualizations for transfer tasks. Moreover, it was expected that spatial abilities would moderate the effectiveness of learning with dynamic and static visualizations. However, dynamic and static visualizations did not differ with regard to any of the learning outcome measures, and spatial abilities also did not moderate learning with these two types of visualizations. In general, higher spatial abilities were associated with higher learning outcomes, indicating that higher spatial abilities are beneficial for learning with dynamic as well as static visualizations (cf. Hegarty & Kriz, 2008; Höffler, 2010). The instructional equality of dynamic and static visualizations may be partly traced back to the fact that learners in the static visualizations condition tended to play the visualizations more often than learners in the dynamic condition. This more frequent use of static visualizations may be interpreted as a strategy to compensate for a drawback of the static presentation format, namely, its demands for mental animation. This interpretation receives further indirect support by the finding that spatial ability moderated the frequency of using static and dynamic visualizations. A similar pattern of results was observed in a study by Schnotz and Rasch (2005; Exp.1), where learners with weaker learning prerequisites spent more time studying static visualizations than studying dynamic visualizations, while the opposite was true for learners with stronger learning prerequisites. The compensatory-strategy interpretation might also account for the finding that spatial ability did not moderate learning with dynamic and static visualizations. Moreover, it would also explain that there were no differences with respect to cognitive load between these two conditions. Hence, it may be the case that in a system-paced learning environment, learning with dynamic visualizations might yield better performance than learning with static visualizations for a dynamic domain like the one at hand. It should be noted though that solely data were collected concerning the frequency of playing the visualizations, but not whether learners actually watched the visualizations. For investigating this issue, for ongoing studies learners' viewing behavior might be recorded by means of eye-tracking data (e.g., Ozcelik, Karakus, Kursun, & Cagiltay, 2009; Scheiter & van Gog, 2009; She & Chen, 2009). Although the compensatory-strategy interpretation is notional, it once again reveals the importance of considering strategic variables, as they may moderate or mediate performance for different instructional conditions (e.g., Gerjets & Scheiter, 2003).

Another reason for the instructional equality of the two visualization conditions may be that the potentials of dynamic visualizations for this domain might not have been exploited completely. Even though high-quality animations were used, one possible benefit of the dynamic visualizations was not used: The interplay of changes in the frequency of the undulatory movement that are associated with changes in the magnitude of the reaction force, and their interrelatedness with swimming speed were not explained in this study. However, the power to depict changes in velocity is an inherent property of dynamic visualizations, which is not available in static frames. To further improve the instructional material, the abovementioned aspects were implemented in the dynamic visualizations used in Study 2 and 3.

Concerning learning strategies, it was assumed that more erroneous as well as more negative monitoring statements would occur in the static visualization condition compared to the dynamic visualization condition. This assumption was confirmed for the number of erroneous statements. Learners in the static visualizations condition may, at least initially, have had difficulties in interpreting the changes of objects depicted in the static visualizations. However, they might not have noticed these difficulties as reflected in the equal amount of negative monitoring statements in the two visualization conditions. In contrast to the results of the study by Lewalter (2003), no differences for rehearsal strategies or for planning for further learning could be observed between the visualization conditions. However, as expected, and in line with Lewalter (2003), learners with dynamic visualizations produced more positive monitoring statements and were therefore more confident that they had understood the content, even though this was not the case as indicated by the learning outcome measures. Moreover, positive monitoring statements did not correlate with performance, suggesting that learners misjudged their understanding. Accordingly, the higher number of positive monitoring statements in the dynamic visualizations condition may be interpreted in terms of an illusion of understanding (cf. Bétrancourt, 2005). It should be noted, however, that this illusion of understanding did not lead to less engagement in valuable processing activities such as elaborations. Therefore, these results cannot be interpreted as evidence that dynamic visualizations – as opposed to a sequence of static visualizations – lead to a shallower processing.

Concerning the relationship between the categories of the think-aloud protocols and learning outcome measures, rehearsal strategies were only associated with verbal factual knowledge. This seems plausible, as rehearsal should mainly support the construction of a verbal model, which might be best assessed by the verbal factual knowledge task. Similar results were also obtained by Lewalter (2003). Surprisingly, producing erroneous statements was not associated with lower learning outcomes. On the one hand, the overall number of erroneous statements can be regarded as being rather low. Moreover, it may be the case that learners producing erroneous statements initially had the opportunity to correct their misunderstanding in the later process of learning, so that these misconceptions may not have had a strong impact. The finding that generating inferences was associated with higher performance for all learning outcomes measures, and that negative monitoring statements were associated with lower performance may be regarded as a positive validation check. However, the overall size of the correlations between the learning strategies and the learning outcome measures are rather low to moderate (Cohen, 1992). It can only be reported anecdotally that some learners had difficulties with the think-aloud procedure, felt uncomfortable with this method and got tired of it (keeping in mind that the average learning time was approximately 24 minutes). Accordingly, these results should be treated with caution. Similar problems with learners applying the think-aloud procedure in learning with visualizations were also reported by Cohen and Hegarty (2007).

Overall, this study can be considered as a contribution to our understanding of how instructional formats can differently influence performance in diverse learning outcome measures, by having shown, for instance, that the multimedia effect only held for pictorial and transfer tasks, but not for verbal factual knowledge tasks. Moreover, it also might pave the way to pay more attention to learning activities in future studies as they may weaken the deterministic relationship that is sometimes assumed between instructional design and learning outcomes (Gerjets & Scheiter, 2003). For instance, whether dynamic visualizations lead to better learning outcomes than static visualizations may depend on the type of learning activities deployed, such as retrieving static visualizations more frequently. Finally, as the current study indicated, conducting think-aloud protocols can be considered a fruitful way to gain deeper insights into the cognitive processes of learners dealing with different instructional materials and their contribution to performance. For instance, it seems to be the case that adding visualizations to text encouraged learners to engage in more valuable processing activities, which in turn helped to achieve a better understanding. The assessment of process data allows for the more direct investigation of specific claims, such as whether visualizations foster inferences, or whether in learning with dynamic visualizations an illusion of understanding might arise, thereby enriching an outcome-oriented approach to investigating multimedia learning. Such a research strategy may help at getting a more thorough understanding of the mechanism underlying learning with multimedia and may be more often considered in future studies in this field.

To conclude, the addition of visualizations to text led to a deeper understanding for this domain. Nonetheless, the multimedia material might be further improved in several ways. First, the redundancy of text and visualizations can be reduced by deleting parts of the text describing visuo-spatial aspects, which are supposed to be more efficiently depicted in the visualizations (cf. Schmidt-Weigand & Scheiter, 2011). This was done for the multimedia material of the following studies in the current thesis. In turn, this also means that learners need to rely more on the visualizations, thereby allowing differences in the instructional efficiency of different visualization formats to become more evident. Moreover, not only visuo-spatial aspects were deleted from the text, but also dynamic features were less extensively described in the text of the following studies. This was done to further reduce the redundancy of dynamic features depicted in dynamic visualizations and dynamic features described in the text. Nevertheless, the dynamic features were not completely eliminated from the text, but only to a degree that they still could be regarded as being comprehensible when learners did not receive dynamic visualizations, but static visualizations.

Second, particularly the design of the dynamic visualizations was rather suboptimal in the current study. They might be further enhanced by depicting dynamic features that can be represented in dynamic visualizations, but not in static visualizations. More precisely, the interplay of changes in the frequency of the undulatory movement that are associated with changes in the magnitude of the reaction force (symbolized by the length of arrows), and their interrelatedness with swimming speed were not depicted in Study 1, but they will be depicted in Study 2 and 3. Specifically, for representing the swimming speed, landscape background has to be added, so that a reference point is given to learners that allows in inferring how much distance the fish had travelled. However, it should be noted that adding background may also increase the visual complexity of the visualizations, and specifically of the dynamic visualizations, because it increases the chance that some less relevant details from the background can become more salient through their movement. Thereby, because more elements are moving at the same time, this may intensify the intra-representation split-attention effect. Due to the assumed high degree of visual complexity in dynamic visualizations, it may be the case that their benefits to depict the dynamic interrelations of a domain like the one at hand might not completely unfold. To improve learning with visualizations in general, and dynamic visualizations in particular, it may be necessary to overcome problems associated with a high degree of visual complexity. These problems are on the one hand an inter-representation split-attention effect between text and visualizations, and on the other hand the visual complexity of visualizations themselves. Therefore, two design characteristics that aim at dealing with these two potential drawbacks will be introduced next, namely using spoken text to overcome inter-representational split attention, as well as cueing to cope with the visual complexity of dynamic visualizations themselves.

4 Optimizing Learning from Visualizations

In the following, first the role of using spoken text to reduce inter-representational split attention, and its potential impact on learning with dynamic as well as static visualizations will be explicated. Thereafter, the role of cueing, as a means to reduce the visual complexity within visualizations will be outlined.

4.1 Reducing Inter-representational Split-attention by Using Spoken Text

The inter-representational split-attention effect occurs when learners have to split their attention between physically separated multiple sources, for instance, if a written text and the corresponding visualization are distributed across pages, but need to be mentally integrated in order to understand the information (cf. Ayres & Sweller, 2005; Sweller et al., 1998). Thereby, inter-representational split-attention is supposed to hamper learning. In the current thesis, a focus lies on avoiding this problem by using spoken text.

Advantages of using spoken text rather than written text to accompany visualizations are referred to as the modality effect in multimedia learning. As with the multimedia effect, the modality effect could be confirmed in many studies (cf. Ginns, 2005; Low & Sweller, 2005; Mayer, 2001, 2009; Sweller et al., 1998). In a meta-analysis by Ginns (2005), a moderate to large overall advantage (d = .72) of spoken over written text for multimedia material could be observed.

The explanation for how spoken text may reduce inter-representational split-attention constitutes the most prominent explanation of the modality effect, namely, the *split-attention explanation* (Rummer, Schweppe, Fürstenberg, Seufert, & Brünken, 2010; see also temporal and spatial contiguity explanation, respectively, Rummer, Schweppe, Fürstenberg, Scheiter, & Zindler, in press; Schüler, Scheiter, Rummer, & Gerjets, 2011)[11]. Note that the split-attention explanation of the modality effect is formulated on the basis of the split-attention effect described by the CLT, which states[12]: "The modality effect derives from the split-attention effect. It occurs under split-attention conditions when a written source of information that must be integrated with another

[11] An alternative to the split-attention explanation of the modality effect is the visuo-spatial load explanation (cf. Rummer et al., 2010). According to this explanation, written text and visualizations compete for limited visuo-spatial working memory resources, whereas spoken text and visualizations can be processed in different subsystems of working memory. However, the underlying theoretical rationale of the modality effect has been challenged (Rummer et al., in press; Rummer et al., 2010; Rummer, Schweppe, Scheiter, & Gerjets, 2008; Schüler et al., 2011; see also Tabbers, 2002). Moreover, a recently conducted study by Rummer et al. (in press) revealed no evidence for the visuo-spatial load explanation.

[12] While in the terminology of CLT it is solely termed split-attention, in the remainder of this chapter it will be referred to as inter-representational split-attention to distinguish it from intra-representational split-attention.

source of visually presented information such as a diagram is presented in auditory rather than visual (written) mode" (Sweller et al., 1998, p. 282). Accordingly, presenting visualizations with spoken text may offer at least three advantages with regard to reducing inter-representational split-attention. First, with spoken text, text and visualizations can be processed simultaneously. Hence, sufficient attention can be devoted to both representations. Second, since spoken text and visualizations are simultaneously available, this might facilitate the integration of text and visualizations. This is assumed to be the case, because with written text, the text has to be held and possibly be reconstructed in working memory before it can be integrated with the visualization (and vice versa). On the other hand, when visualizations are accompanied by spoken text, a learner can process both representations simultaneously, without needing to hold the information longer than necessary in working memory (cf. temporal contiguity explanation). Third, in contrast to written text, with spoken text, no switches have to be made between text and visualizations, thereby reducing visual search demands. Furthermore, as with written text and visualizations the visual attention has to be shifted between the two representations, this is supposed to pose additional perceptual demands on the visual system (cf. spatial contiguity explanation). In the terminology of CLT, the latter two advantages are usually associated with a decrease of ECL when using spoken text. Support for the assumption that spoken text reduces ECL is given by several studies, which assessed either subjective cognitive load ratings (e.g., Kalyuga, Chandler, & Sweller, 2000; Tindall-Ford, Chandler, & Sweller, 1997; van Gerven, Paas, van Merriënboer, & Schmidt, 2006), or used a dual task methodology (e.g., Brünken, Steinbacher, Plass, & Leutner, 2002).

Consequences of using spoken text to reduce inter-representational split-attention may differ as a function of the mental models that emerge during multimedia learning. In particular, Schmidt-Weigand et al. (2010) showed that under inter-representational split-attention conditions, written text was processed superordinate (see also Hegarty & Just, 1993), resulting in less time allocated to the visualizations, but possibly in sufficient time allocated to the text. Accordingly, one might assume that learners receiving written text and visualizations will develop a comprehensive verbal mental model, but compared to learners receiving spoken text and visualizations, a less comprehensive pictorial mental model and, in connection, also a less developed integrated mental model. As a consequence, the modality effect should be more pronounced for pictorial tasks as well as transfer tasks as opposed to verbal factual knowledge tasks. First evidence for this assumption was reported in the meta-analysis by Ginns (2005), where the modality effect emerged for transfer tasks, but not for verbal factual knowledge tasks. Furthermore, there are indications that the modality effect also holds for pictorial tasks (e.g., Craig, Gholson, & Driscoll, 2002; Mayer & Moreno, 1998; Moreno & Mayer, 1999; Rummer et al., in press; Schmidt-Weigand et al., 2010; Schüler et al., 2011), even though these kind of tasks are rarely assessed in current research on multimedia learning.

Overall, spoken text might be a remedy to cope with the problems arising from an inter-representational split-attention effect. The inter-representational split-attention effect may be especially problematic, if the visual complexity of the visualizations is high, which is assumed to be the case specifically for dynamic visualizations (cf. Chapter 2.2.2). Thus, when combining dynamic visualizations with written rather than spoken text, this may further increase the visual search demands, and, accordingly, may make it particularly hard to process the information conveyed by dynamic visualizations. Hence, the potentials of dynamic visualizations may not properly unfold with written text. On the other hand, when presenting spoken text, learners do not have to split their attention between text and visualization. In this case, the information in the dynamic visualizations may be selected and extracted appropriately, so that their potentials may be able to unfold. Therefore, the advantages of dynamic compared to static visualizations are expected to be more pronounced with spoken text. Schnotz (2005) arrives at similar conclusions when stating: "The negative effects of split-attention on learning are especially pronounced when animated pictures are used instead of static pictures" (p. 61), and "split-attention becomes less important if static pictures are used" (p. 65).

Empirical evidence

Support for the claim that the superiority of dynamic visualizations as opposed to static visualizations may be more pronounced for spoken than for written text may come from the meta-analysis of Höffler and Leutner (2007), the conducted studies since this meta-analysis, as well as at the meta-analysis of Ginns (2005) on the modality effect in multimedia learning[13].

While Höffler and Leutner (2007) considered the presence and absence of text as a potential moderator in learning with dynamic and static visualizations, they neglected the modality of the text. To examine the assumption that spoken text might particularly be beneficial for learning with dynamic visualizations, the meta-analysis by Höffler and Leutner (2007) was scrutinized. In their meta-analysis, the authors identified 56 comparisons from 20 studies that included text at all[14]. By reinspecting these 56 comparisons, several studies and the comparisons contained in them had to be excluded. First, the study by Rigney and Lutz (1976) did not compare dynamic visualizations to static visualizations. Second, the study by Spotts and Dwyer (1996)

[13] It should be noted that the two meta-analyses differed in their applied methods of how to calculate the effect size, and hence cannot be simply merged for a re-examination.
[14] Höffler and Leutner stated 59 comparisons, but adding up their coding revealed only 56 comparisons.

compared interactive animations to non-interactive static visualizations and hence violated the request of equivalent procedures (cf. Tversky et al., 2002), even though Höffler and Leutner (2007) stated to control for this factor to have “fair” comparisons. Third, for the study by Lai (2000), the comparison of dynamic and static visualizations which included spoken text was excluded, because a negative effect size was reported, even though learners with dynamic visualizations clearly outperformed learners with static visualizations. It should be noted that because of the exclusion of the study by Lai (2000) the mean effect size for the comparison of dynamic versus static visualizations that are accompanied by spoken text will be underestimated. Taken together, this resulted in 17 studies with 49 comparisons, whereof 38 comparisons included written text, and 11 comparisons included spoken text (cf. Table 4.1). The arithmetic means of the effect sizes for the comparisons of dynamic and static visualizations were calculated as a function of the modality of the text. For written text, there was a mean effect size of $d = .25$ ($SD = .70$) in favor of dynamic visualizations, and for spoken text there was a mean effect size of $d = .15$ ($SD = .54$) in favor of dynamic visualizations. A t-test with text modality as between subject factor and effect size as dependent variable showed no significant difference between the superiority of dynamic over static visualizations for written text as compared to the superiority of dynamic over static visualizations for spoken text ($t(47) = 0.43$; $p = .67$). To conclude, with respect to the studies contained in this meta-analysis by Höffler and Leutner (2007), there was no direct support for the assumption that text modality might moderate the effectiveness in learning with dynamic as opposed to static visualizations. As abovementioned, the mean effect size for comparisons building on spoken text is probably underestimated (because of the wrong effect size of the study by Lai, 2000), but on the other hand, even the correct effect size would probably not change the results dramatically. Nevertheless, these results should be treated with caution. Therefore, as a next step the recently conducted studies in learning with dynamic and static visualizations were inspected on a descriptive level and, furthermore, the meta-analysis by Ginns (2005) was reinspected.

Table 4.1

Effect Sizes in Favor of Dynamic Visualizations as Compared to Static Visualizations from Studies of the Meta-analysis by Höffler and Leutner (2007), Listed as a Function of Text Modality

	Authors	Text Modality	Weighted Effect Size *d*
1	Baek & Layne, 1988	Written Text	.58
2	Baek & Layne, 1988	Written Text	.35
3	Catrambone & Seay, 2002 (Exp.2)	Written Text	.74
4	Catrambone & Seay, 2002 (Exp.2)	Written Text	-1.19
5	ChanLin, 1998	Written Text	.37
6	ChanLin, 1998	Written Text	-.38
7	ChanLin, 1998	Written Text	0
8	ChanLin, 1998	Written Text	.11
9	ChanLin, 2001	Written Text	1.13
10	ChanLin, 2001	Written Text	.97
11	ChanLin, 2001	Written Text	-.86
12	ChanLin, 2001	Written Text	-1.13
13	Hays, 1996	Written Text	.11
14	Hays, 1996	Written Text	1.15
15	Lewalter, 2003	Written Text	0
16	Lewalter, 2003	Written Text	.18
17	Nerdel, 2003 (Exp. 2)	Written Text	.3
18	Nerdel, 2003 (Exp. 2)	Written Text	.38
19	Nerdel, 2003 (Exp. 3)	Written Text	.04
20	Nerdel, 2003 (Exp. 3)	Written Text	-.2
21	Nicholls, & Merkel, 1996 (Exp. 1)	Written Text	.27
22	Rieber, 1989	Written Text	.08
23	Rieber, 1989	Written Text	.11
24	Rieber, 1989	Written Text	.04
25	Rieber, 1989	Written Text	.07
26	Rieber, 1989	Written Text	-.01
27	Rieber, 1989	Written Text	-.11
28	Rieber, 1989	Written Text	-.09
29	Rieber, 1989	Written Text	-.01
30	Rieber, 1990	Written Text	1.32
31	Rieber, 1991	Written Text	.73
32	Rieber, 1991	Written Text	1.48
33	Rieber, Boyce, & Assad, 1990	Written Text	.04
34	Rieber, Boyce, & Assad, 1990	Written Text	-.21
35	Rieber, Boyce, & Assad, 1990	Written Text	.22
36	Szabo & Poohkay, 1996	Written Text	2.63
37	Wright, Milroy, & Lickorish, 1999	Written Text	.24
38	Wright, Milroy, & Lickorish, 1999	Written Text	-.09
39	Craig et al., 2002 (Exp. 1)	Spoken Text	.01
40	Craig et al., 2002 (Exp. 1)	Spoken Text	.06
41	Craig et al., 2002 (Exp. 1)	Spoken Text	-.01
42	Craig et al., 2002 (Exp. 1)	Spoken Text	.04
43	Craig et al., 2002 (Exp. 1)	Spoken Text	.02
44	Craig et al., 2002 (Exp. 1)	Spoken Text	.08
45	Craig et al., 2002 (Exp. 1)	Spoken Text	-.01
46	Craig et al., 2002 (Exp. 1)	Spoken Text	.03
47	Höffler, 2003	Spoken Text	-.19
48	Höffler, 2003	Spoken Text	-.15
49	Yang et al., 2003	Spoken Text	1.76

From the recently conducted 34 studies that include 42 experiments on learning with dynamic and static visualizations, there were 28 experiments where a comparison of dynamic and static visualizations with respect to the text modality was retraceable[15]. Of these 28 experiments, 15 experiments included written text, and 13 included spoken text (see Table 4.2). Of the 15 experiments that included written text, seven showed a positive effect in favor of dynamic visualizations (46.67%), whereas eight showed no differences, or even an advantage for static visualizations (53.33%). Of the 13 experiments that included spoken text, eleven showed a positive effect in favor of dynamic visualizations (86.67%), whereas two experiments showed no differences (13.33%). Fisher's exact test (one-tailed) revealed that this relationship was marginal significant (p = .09)[16]. When only considering experiments that are on the one hand methodological sound and, moreover, where every comparison of dynamic and static visualizations led to a superiority of dynamic visualizations, only 17 experiments can be taken into account. Of these 17 experiments, seven included written text, whereby three of them (42.86%) showed a superiority of dynamic over static visualizations, and ten included spoken text, whereby eight of them showed a superiority of dynamic visualizations (80.00%). On a descriptive level this finding mirrors the pattern of results identified for all 28 experiments, even though Fisher's exact test (one-tailed) failed to reach statistical significance (p = .15), which in turn also may be considered as a power problem. Nevertheless, broadly speaking, this finding may be cautiously interpreted as suggesting that the superiority of dynamic over static visualizations is more pronounced for spoken than for written text.

[15] In a strict sense, these were only 27 experiments. However, since the study of Koroghlanian and Klein (2004) incorporates two comparisons, one for written and one for spoken text, this study was treated as two studies, leading to a count of 28 experiments in Table 4.2.
[16] Fisher's exact test was used, since the preconditions for using a chi-square test were not given.

Table 4.2

Effects in Favor of Dynamic Visualizations From Studies Comparing Learning With Dynamic and Static Visualizations Published Since 2004, Listed as a Function of Text Modality

	Authors	Text Modality	Effect for Dynamic Visualizations
1	Boucheix & Guignard, 2005	Written text	Positive
2	Boucheix & Schneider, 2009 [b]	Written text	Positive
3	Iskander & Curtis, 2005 [a]	Written text	Positive
4	Lin & Dwyer, 2010	Written text	Positive
5	Lin et al., 2009	Written text	Positive
6	Schnotz & Rasch, 2005 [a]	Written text	Positive
7	Wang et al., 2011 [a]	Written text	Positive
8	Yarden & Yarden, 2010 [a]	Written text	Positive
9	Höffler et al., 2010	Written text	Neutral
10	Tunuguntla et al., 2008 [a]	Written text	Neutral
11	van Oostendorp & Beijersbergen, 2007	Written text	Neutral
12	van Oostendorp et al., 2008 [a]	Written text	Neutral
13	Zhu & Grabowski, 2006 [a]	Written text	Neutral
14	Scheiter et al., 2006	Written text	Negative
15	Koroghlanian & Klein, 2004	Written text	Neutral
16	Koroghlanian & Klein, 2004	Spoken text	Neutral
17	Höffler, 2007 (Exp. 3)	Spoken text	Neutral
18	Arguel & Jamet, 2009	Spoken text	Positive
19	Höffler, 2007 (Exp. 1)	Spoken text	Positive
20	Höffler, 2007 (Exp. 2)	Spoken text	Positive
21	Imhof et al., 2009 [b]	Spoken text	Positive
22	Imhof et al., 2010	Spoken text	Positive
23	Kim et al., 2007 [b]	Spoken text	Positive
24	Münzer et al., 2009 [b]	Spoken text	Positive
25	Pfeiffer et al., 2009	Spoken text	Positive
26	Rebetez et al., 2010	Spoken text	Positive
27	Stebner, 2009	Spoken text	Positive
28	Wong et al., 2009 (Exp. 1)	Spoken text	Positive

[a] Note. Studies that are not methodological sound, for instance in terms of a “fair” comparison as recommended by Tversky et al. (2002).
[b] Note. These studies included either different types of dynamic visualizations or different types of static visualizations, whereas not every comparison was in favor of dynamic visualizations.

As previously mentioned, in the meta-analysis by Ginns (2005) several potential moderating variables were accounted for, such as the kind of knowledge test. However, the type of visualization was neglected. Therefore, the 39 cited between-subject experiments of this meta-analysis were reinspected with respect to the question of whether the visualizations were dynamic or static[17]. By doing so, two experiments by Levin and Divine-Hawkins (1974) had to be excluded, because these did not show visualizations in the learning phase, but instructed learners to create mental images instead. From the remaining 37 experiments, 16 used dynamic visualizations, and 21 used static visualizations (cf. Table 4.3). The arithmetic means of the effect

[17] The four within-subject experiments were excluded as they are also analyzed separately by Ginns. Even if they had been considered in the re-examination, this would not have considerably changed the results.

sizes for the respective types of visualizations were calculated. For dynamic visualizations there was a mean effect size of *d* = 1.13 (*SD* = .57) in favor of spoken text, whereas for static visualizations there was a mean effect size of *d* = .50 (*SD* = .60). A t-test with type of visualization as between subject factor and effect size as dependent variable showed a significant effect ($t(35) = 3.19$; $p < .01$) between the benefits of spoken text for dynamic visualizations as opposed to the benefits of spoken text for static visualizations.

Table 4.3

Effect Size in Favor of Spoken Text from Studies from the Meta-analysis by Ginns (2005), Listed as Function of Type of Visualization

	Authors	Type of Visualization	Effect Size *d*
1	Atkinson, 2002 (Exp. 1)	Dynamic	.82
2	Atkinson, 2002 (Exp. 2)	Dynamic	.35
3	Craig et al., 2002 (Exp. 1)	Dynamic	.93
4	Mayer & Moreno, 1998 (Exp. 1)	Dynamic	.46
5	Mayer & Moreno, 1998 (Exp. 2)	Dynamic	.76
6	Mayer, Dow, & Mayer, 2003 (Exp. 1)	Dynamic	.78
7	Moreno & Mayer, 1999 (Exp. 1)	Dynamic	1.49
8	Moreno & Mayer, 1999 (Exp. 2)	Dynamic	1.13
9	Moreno & Mayer, 2002 (Exp. 1)	Dynamic	2.51
10	Moreno & Mayer, 2002 (Exp. 1)	Dynamic	.91
11	Moreno & Mayer, 2002 (Exp. 2)	Dynamic	2.12
12	Moreno & Mayer, 2002 (Exp. 2)	Dynamic	.61
13	Moreno & Mayer, 2002 (Exp. 1)	Dynamic	.60
14	Moreno, Mayer, Spires, & Lester, 2001 (Exp. 4)	Dynamic	1.09
15	Moreno, Mayer, Spires, & Lester, 2001 (Exp. 5)	Dynamic	1.54
16	O'Neil, Mayer, Herl, Niemi, Olin, & Thurman, 2000	Dynamic	.97
17	Brünken & Leutner, 2001	Static	.58
18	Jeung, Chandler, & Sweller, 1997 (Exp. 3)	Static	1.03
19	Kalyuga, Chandler, & Sweller, 1999	Static	1.32
20	Leahy, Chandler, & Sweller, 2002 (Exp. 1)	Static	.04
21	Mousavi, Low, & Sweller, 1995 (Exp. 1)	Static	.93
22	Mousavi, Low, & Sweller, 1995 (Exp. 2)	Static	.88
23	Mousavi, Low, & Sweller, 1995 (Exp. 3)	Static	.65
24	Mousavi, Low, & Sweller, 1995 (Exp. 4)	Static	.68
25	Mousavi, Low, & Sweller, 1995 (Exp. 5)	Static	.63
26	Tabbers, 2002 (Chapter 4)	Static	.51
27	Tabbers, 2002 (Chapter 4)	Static	.09
28	Tabbers, 2002 (Chapter 4)	Static	-.66
29	Tabbers, Martens, & van Merriënboer, 2000	Static	-.54
30	Tabbers, Martens, & van Merriënboer, 2001 (Exp. 1)	Static	.67
31	Tabbers, Martens, & van Merriënboer, 2001 (Exp. 2)	Static	-.06
32	Tabbers, Martens, & van Merriënboer, 2001 (Exp. 2)	Static	.73
33	Tabbers, Martens, & van Merriënboer, 2004	Static	-.47
34	Tindal-Ford et al.,1997 (Exp. 1)	Static	1.68
35	Tindal-Ford et al.,1997 (Exp. 2)	Static	1.07
36	Tindal-Ford et al.,1997 (Exp. 3)	Static	.23
37	van Gerven (Chapter 5), 2002	Static	.59

Accordingly, reinspecting the meta-analysis of Ginns (2005) revealed that the modality effect was more pronounced for dynamic visualizations than for static visualizations. This also indicates that dynamic visualizations as opposed to static visualizations might suffer more from written text or profit more from spoken text, respectively. Hence, this re-examination can be seen as a further hint towards the assumption that text modality might moderate learning with dynamic and static visualizations – even though such results from meta-analyses should be treated with caution, because there mostly will be further confounding variables that cannot all be reasonably taken into account.

There are only two published studies to the author’s knowledge that investigated the role of text modality in learning with dynamic and static visualizations (Koroghlanian & Klein, 2004; Mayer et al., 2005). Koroghlanian and Klein (2004) observed neither a main effect for type of visualizations, nor a main effect for text modality, nor an interaction between these two factors. However, in this study, the spoken text conditions additionally received a stripped-down version of the same text in written form next to the visualizations, so that the potential benefit of spoken text, namely to reduce inter-representational split-attention, might have been non-existent. Moreover, presenting spoken text and a stripped-down version of written text together might have even led to a harmful redundancy, thereby further eliminating the potential benefit of spoken text in multimedia learning (cf. redundancy principle, Sweller, 2005b). In every experiment of the study by Mayer et al. (2005), solely two conditions were compared, namely a condition with spoken text and dynamic visualization to a condition with written text and static pictures. Due to the confounded nature of the chosen design of this study, however, the role of text modality in learning with dynamic and static visualizations could not be examined.

To sum up, dynamic visualizations possess high potentials for conveying a deeper understanding of changes over time. However, due to their assumed visual complexity, these potentials may not properly unfold under inter-representational split-attention conditions. To cope with this problem, it might be reasonable to use spoken text. When optimizing learning with dynamic visualizations by reducing inter-representational split-attention through using spoken text, it may be the case that the superiority of dynamic visualizations as opposed to static visualizations might be even more pronounced. Since there appears to be no published study that directly addressed this question, this question was examined in Study 2.

It should be noted though that using spoken text does not reduce the visual complexity of dynamic visualizations themselves. To counteract the processing demands associated with a high degree of visual complexity, it has been suggested to cue visualizations. This topic will be addressed next.

4.2 Reducing Visual Complexity by Cueing Important Information

When considering learning with dynamic visualizations, learners may be overwhelmed by the high degree of visual complexity within dynamic visualizations (e.g., Lowe, 2003, 2004). This high degree of visual complexity might be induced by at least three factors: By the continuous movement of elements, by the distracting movement of less relevant elements, and by the need to split attention within dynamic visualizations since several elements may move at different locations at the same time. One way to overcome the processing demands of dynamic visualizations that are associated with their visual complexity is assumed to lie in the use of cueing techniques (cf. Bétrancourt, 2005; Boucheix & Lowe, 2010; de Koning, Tabbers, Rikers, & Paas, 2009; Schnotz & Lowe, 2008).

In line with de Koning et al. (2009), cueing is referred to as “the manipulation of visuo-spatial characteristics of instructional material in order to help learners in selecting relevant information, and organizing and integrating the information into a coherent representation” (p. 114). Thereby, by means of cueing, no additional information is added.

According to Mayer (cf. signaling principle, 2005c, 2009), learning is facilitated when the essential information of the instructional material is cued. Even though Mayer (2005c, 2009) refers to the respective effect as evidence for the signaling principle, in the following the term cueing will be used as it is most often used in current multimedia research (see also de Koning et al., 2009). On the basis of the CTML, cueing should aid learning because it guides the learners’ attention towards the essential information thereby leaving more resources available for more thorough processing of the essential material (Mayer, 2005c, 2009; Mayer & Moreno, 2003). In CLT terminology, this corresponds with a decrease of ECL so that more resources may be devoted to processes associated with GCL.

The potential effects of cueing can be well described by referring to the processes of selection, organization and integration within the CTML as will be shown in the following (cf. de Koning et al., 2009, Mautone & Mayer, 2001; Mayer & Moreno, 2003)[18]. Thereby, most of the cueing methods that support the processes of selection, organization, and integration can also be related to their potential to reduce the visual complexity within visualizations[19].

[18] Surprisingly, Mayer (2001, 2005c, 2009) provides hardly any explanation of cueing within the CTML with regard to the processes of selection, organization, and integration, and the emerging mental models in his textbooks.

[19] However, it should be noted though that not every type of cue that support one of these processes necessarily reduces visual complexity. Therefore, in the reminder of this chapter, these different functions will also be discussed with regard to the type of cue.

First, cueing may help to *select* relevant information. Concerning text, the selection of important words may be facilitated by means of cueing, for instance by using bold type (in the case of written text) or by a different intonation (in the case of spoken text), which is supposed to make these words more distinguishable from other text elements. Concerning visualizations, the selection of visual information might be disburdened by cueing essential elements within the visualization, for instance, by means of coloring these elements or by using a spotlight. It is assumed that cueing aids the selection of relevant information because on the one hand, by means of cueing, attention can be guided. Here, it also may facilitate the discrimination process that singles out important information, thereby enabling a more intensive processing of that information (cf. Mautone & Mayer, 2001). Recent research applying eyetracking methods supports the assumption of an attention-guiding function of cueing (e.g., Boucheix & Lowe, 2010; de Koning et al., 2010a; Jarodzka, 2011; Kriz & Hegarty, 2007; Ozcelik, Arslan-Ari, & Cagiltay, 2010). For instance, de Koning et al. (2010a) used a spotlight that successively highlighted several subsystems of an animation depicting the cardiovascular system. Eyetracking data revealed that learners looked more often and longer at the cued region as opposed to learners in a control condition. Moreover, cueing may reduce unnecessary search processes so that “less visuospatial recourses are required to control the execution of eye movements. Thereby, cueing reduces extraneous cognitive load associated with locating relevant information.” (de Koning et al., 2009, p. 118). By reducing visual search processes, the demands on the perceptual system might be disburdened. First evidence for the claim that cueing additionally might reduce unnecessary search processes is also given by recent eyetracking studies (Jarodzka, 2011; Exp. 2; Ozcelik et al., 2010). On the other hand, the eyetracking data from de Koning et al. (2010a) as well as Kriz and Hegarty (2007) did not reveal a reduction of visual search. However, as Ozcelik et al. (2010) points out, this may be due to inadequate analysis of the eyetracking data as, for instance, de Koning et al. (2010a) did not use a time-locked analysis (cf. Hyönä, 2010). Irrespective of this, by guiding a learner’s attention and by reducing visual search processes within a visualization, cueing (specifically a spotlight cue) might reduce problems associated with the visual complexity of visualizations, particularly the problems of distracting information, as well as intra-representational split-attention.

Cueing may also help in *organizing* the information. Thereby, text might be cued by means of headlines, enumerations, or summaries, which are supposed to make the global and local structures of a text more evident (e.g., Loman & Mayer, 1983; Lorch & Lorch, 1996; Meyer, 1975; Mayer & Moreno, 2003). This in turn may help learners to organize the material more easily into a coherent verbal model (Mautone & Mayer, 2001). With respect to organizing information in visualizations by means of cueing, one might, for instance, cue the relevant elements one after the other to stress the functional order of a system, or to mirror the cause-and-effect chain, respectively (e.g., Boucheix & Lowe, 2010). Thereby, cueing may not only help to organize a visualization in terms of its spatial order, but also with respect to the temporal order of the events depicted. Here, the processing of the respective information is guided, and, in connection, the visual complexity might be further reduced, since learners might be less distracted from information that is less relevant at a given point in time. Also, elements belonging together could, for instance, be grouped by depicting the same color, since they belong to the same functional unit within a system (cf. de Koning et al., 2009). Doing so might decrease ECL and might make it easier for the learner to organize the elements within a visualization into a coherent pictorial model. Moreover, by guiding a learner's processing, cueing might additionally help the learner to better concentrate on the content and to devote the freed capacity to valuable processing activities, which in turn would correspond to an increase in GCL. Summing up, cueing is assumed to help in organizing the content, because the order of processing events is guided, thereby also possibly reducing the visual complexity within visualizations. Moreover, relations can be explicated that otherwise might have to be inferred – a process that can be considered resource-demanding (cf. de Koning et al., 2009; Mautone & Mayer, 2001, 2007). Since specifically the use of effective visual cues for visualizations should result in a more elaborated pictorial model, the effectiveness of cueing might result in better performance on pictorial tasks. In line with this assumption, a positive effect of cueing on pictorial tasks could be observed (e.g., Beck, 1987; Boucheix & Guignard, 2005; Ozcelik et al., 2010; Van Meter, Gu, Pastore, & Cook, 2010), even though it should be noted that this learning outcome measure is seldom assessed in research on cueing.

Cueing might not only help to select and organize the information *within* a representation, but also to relate information *between* two (or more) representations. Thereby, cueing might aid the *integration* of information of different external representations into a coherent mental representation of the content (cf. de Koning et al., 2009). This in turn might also be associated with an increase of GCL. Furthermore, ECL might be reduced, since working memory might be freed up by the lack of necessity for the resource-demanding processes of inferring relations and correspondences between elements of text. In the case of written text and visualizations, the integration function of cueing could be realized for instance by means of color-coding, that is by giving the same color to the information in the written text and the referring information in the visualization (e.g., Folker, Ritter, & Sichelschmidt, 2005; Kalyuga et al., 1999; Ozcelik et al., 2009). In case of spoken text and visualizations, the integration function of cueing could be realized by synchronizing the spoken text and the visualizations, for instance, by highlighting elements when mentioned in the narration, or by not adding elements in the visualizations until mentioned in the text (e.g., Jamet, Gavota, & Quaireau, 2008; Jeung et al., 1997; Ozcelik et al., 2010). Thereby, problems associated with intra-representational split-attention might additionally be decreased, since the visualizations gradually build up. Overall, emphasizing relationships between elements of the visualizations and corresponding elements of the text might lead to a well developed integrated mental model, which in turn should result in better performance on transfer tasks. Partly in line with this reasoning, a positive effect of cueing could be observed for transfer tasks (e.g., de Koning, Tabbers, Rikers, & Paas, 2010b; Ozcelik et al., 2010), admittedly not always (e.g., Huk, Steinke, & Floto, 2010; Jamet et al., 2008).

In short, cueing may support the processes of selecting and organizing information, which in turn leads to a more coherent pictorial model when cueing visualizations[20]. Moreover, by supporting learners in relating corresponding verbal and pictorial information to each other, cueing may also facilitate the construction of a well developed integrated model. With regard to the demands on working memory, cueing might reduce ECL and increase GCL. However, there is little direct empirical evidence that cueing indeed reduces ECL as measured by subjective cognitive load ratings (e.g., Amadieu, Mariné, & Laimay, 2011; Berthold & Renkl, 2009; Jamet et al., 2008; Kalyuga et al., 1999). In other studies, no differences in the assessed subjective cognitive load ratings could be observed (e.g., de Koning, Tabbers, Rikers, & Paas, 2007; Keller, Gerjets, Scheiter, & Garsoffky, 2006; Tabbers et al., 2004), even though in the latter studies cueing had an effect on performance (cf. de Koning et al., 2009). Finally, most of the cueing methods that support the processes of selecting, organizing and integrating are also assumed to reduce the visual complexity of visualizations. However, since particularly dynamic visualizations, compared to static visualizations, are supposed to suffer from a high degree of visual complexity, the benefits of cueing might be more pronounced for dynamic than for static visualizations.

Empirical evidence

To pursue the claim that the benefits of cueing might be more pronounced for dynamic than for static visualizations, in the following, the research on the effects of cueing in learning with dynamic and static visualizations will be inspected. Unlike for the influence of using spoken text to optimize learning from dynamic as compared to static visualizations, where the meta-analyses of Höffler and Leutner (2007), Ginns (2005), as well as more recently conducted studies could be re-inspected for estimating this effect, such a procedure could not be applied to determine whether cueing is especially suited to foster learning from dynamic visualizations. This is due to the fact that only a low number of studies covered by the meta-analyses of Höffler and Leutner (2007)

[20] Note that cueing text would similarly lead to a more coherent verbal model. However, as text was not cued in the studies of the current thesis, this topic will be neglected in the following.

incorporated cueing in learning from dynamic and static visualizations (at least as can be gathered from the description of these studies). Therefore, it was not possible to reasonably compare the studies that incorporated cueing in dynamic and static visualizations to those that did not use cueing in dynamic and static visualizations[21]. The same conclusion accounts for the recent studies that compared dynamic and static visualizations, of which only three incorporated cueing in their design for both types of visualizations (Ardac & Akaygun, 2005; Boucheix & Guignard, 2005; Pfeiffer et al., 2009). Finally, no meta-analysis exists with respect to cueing. Based on the information illustrated above, an overview of cueing in static visualizations will first be given. This will be rather brief as the research results regarding the effectiveness of cueing for static visualizations are consistent and mainly positive. Subsequently, the existing research concerning cueing in dynamic visualizations will be considered. This will be done in more detail due to the mixed pattern of results of a recent review by de Koning et al. (2009) considering the research on the effectiveness of cueing in dynamic visualizations. Studies unsuccessfully applying cues in dynamic visualizations will be explicated followed by studies that successfully applied cues in dynamic visualizations, after which suggestions as to what might have caused this different pattern of results for cueing in dynamic visualizations will follow. Finally, conclusions concerning the role of cueing in learning with dynamic and static visualizations will be reached.

Cueing in static visualizations

Research on the effectiveness of cueing in static visualizations is quite consistent and basically shows a positive effect of cueing on performance. In contrast to the research with dynamic visualizations discussed below, the studies entailing cueing in static visualizations mainly used multimedia material (i.e., text and visualizations), with the exception of one study by Grant and Spivey (2003), who solely cued a static visualization and also found a positive effect of cueing. The remaining studies investigated the influence of static visualizations accompanied by text, and the majority of these studies showed a positive effect of cueing on performance. This was the case for studies presenting static visualizations and written text (e.g., Beck, 1984, 1985, 1987; Berthold & Renkl, 2009; Florax & Plötzner, 2010; Folker et al., 2005; Kalyuga et al., 1999; Keller et al., 2006; Mautone & Mayer, 2007; Ozcelik et al., 2009; Scheiter & Eitel, 2010; Seufert & Brünken, 2006; Seufert, Jänen, & Brünken, 2007; Tabbers et al., 2004), as well as for studies presenting static visualizations and spoken text (e.g., Jamet et al., 2008; Jeung et al., 1997; Ozcelik, et al., 2010; Tabbers et al., 2004). Moreover, it could be shown that cueing not only has a positive effect on

[21] It should be noted though that it was possible in the meta-analysis by Höffler and Leutner (2007) to compare dynamic visualizations to static visualizations that contained cues (particularly arrows) and to static visualizations that did not contain cues; there were no differences of cued and uncued static visualizations with regard to their instructional effectiveness.

the comprehension of cued information, but that cueing also does not have a negative effect on the comprehension of uncued information (e.g., Beck, 1985; de Koning et al., 2007). However, it should be noted that some boundary conditions for the effectiveness of cueing emerged: For instance, ICL should not be too high or too low, and similarly, ECL should not be too low, as otherwise the effectiveness of cueing might not unfold (Jeung et al., 1997; Seufert et al., 2007).

As explicated in Chapter 2.2.2, with respect to the type of multiple static visualizations, one can differentiate between static-sequential and static-simultaneous visualizations. However, since in mainly all studies single static visualizations were cued, there is a hardly any research with respect to the effectiveness of cueing in *multiple* static visualizations. From a theoretical point of view, one may nevertheless derive assumptions concerning the effectiveness of cueing in static-sequential as opposed to static-simultaneous visualizations. As mentioned in Chapter 2.2.2, one potential drawback of static-simultaneous as opposed to static-sequential visualizations is that learners have to conduct more visual search and matching processes. Thus, one may assume that learners might particularly need guidance in static-simultaneous visualizations, which can be realized by means of cueing. If this assumption was true, one would suppose that learners receiving static-simultaneous visualizations may profit more from cueing than learners receiving static-sequential visualizations. However, it should be noted that there is a lack of research with respect to this research question, so this assumption stands on shaky ground.

In the following, research concerning cueing in dynamic visualizations will be outlined, beginning with research where no evidence in favor of cueing was observable. An overview of the studies using cues in animations is provided in Table 4.4.

Table 4.4

Overview of Studies Investigating the Effectiveness of Cues in Dynamic Visualizations, Listed as a Function of Text Presence and Modality

Authors	Text modality	Effect of cueing
de Koning et al., 2010a	No text	Neutral
de Koning et al. 2011a	No text	Neutral
Kriz & Hegarty, 2007	No text	Neutral
Mautone & Mayer, 2001	Spoken text	Neutral
Moreno, 2007	Spoken text	Neutral
Spangenberg, 1973	Spoken text	Neutral
Large et al., 1996	Written text	Neutral
van Oostendorp & Beijersbergen, 2007	Written text	Neutral
de Koning et al., 2007	No text	Positive
Meyer et al., 2010	No text	Positive
Boucheix & Lowe, 2010	No text	Positive (but not for arrows)
Fischer, 2008 (Exp. 2)	No text	Positive (for fast speed)
Fischer et al., 2008	No text	Positive (for fast speed)
Fischer & Schwan, 2010	No text	Positive (for fast speed)
de Koning et al., 2011b	No text	Positive (in combination with self-explanations)
Amadieu et al., 2011	Spoken text	Positive
de Koning et al., 2010b	Spoken text	Positive
Huk, 2010	Spoken text	Positive
Huk, 2010	Spoken text	Positive
Jarodzka, 2011 (Exp. 2)	Spoken text	Positive
Jarodzka, 2011 (Exp. 3)	Spoken text	Positive
Janelle et al., 2003	Spoken text	Positive (for multiple cues)
Boucheix & Guignard, 2005	Written text	Positive
van Oostendorp et al., 2008	Written text	Positive

Cueing in dynamic visualizations: no effects

There are several studies that could not find a positive effect of cueing on comprehension of dynamic visualizations. These studies will be described next, beginning with studies using dynamic visualizations without text, followed by studies that were accompanied by written or spoken text, respectively. Thereby, possible explanations will be given for what might have caused the lack of differences in favor of cueing in these studies.

In a series of five studies, de Koning, Tabbers, Rikers, and Paas (2007, 2010a, 2010b, 2011a, 2011b) – with the exception of one study (de Koning et al., 2007) – could not find a superiority of cued compared to uncued animations, unless the animations were supported by further explanations. In these studies, comprehension was unaffected by cueing, even though an analysis of eyetracking data revealed that the cues served to guide attention towards relevant information (de Koning et al., 2010a). Similarly, Kriz and Hegarty (2007; Exp. 3) could show with the help of eyetracking data that cueing guided attention to the cued parts of the animation, even though no differences could be found for comprehension (Exp. 2 & 3). According to de Koning et al. (2010a), these results might be interpreted by suggesting that cueing mainly stimulated perceptual processing rather than cognitive processing, and that cues are seldom efficient to foster understanding as long as no further explanations are given, for instance, by an accompanying text. Similarly, Kriz and Hegarty (2007) stressed the importance "to make a distinction between the perceptual processes of extracting the visual features of a display and the more conceptual processes of encoding that display and constructing a mental model of the referent" (p. 925).

Irrespective of that conclusion, it should be noted though that a drawback in the experiments by Kriz and Hegarty (2007) was that they used arrows as a cueing device. However, arrows may interfere with the depicted motion in an animation, as arrows are not only used to point to specific regions, but also to show the direction of a movement (Heiser & Tversky, 2006). More direct support for the assumption that arrows are a rather suboptimal type of cue for learning with dynamic visualizations is given by a study from Boucheix and Lowe (2010). In this study the authors compared arrows as cues with a spreading-color cue (i.e., a cue that consists of colored ribbons and that spreads through the relevant graphic entities in synchrony with the main causal chain of those entities). The results revealed that only the spreading-color cue, but not arrows were beneficial to enhance understanding. This study and its results will be described in more detail below.

When considering the interplay of several representations, such as animations and text, cues may not only serve to highlight information within a representation, but additionally to relate corresponding elements in the representations to each other (cf. de Koning et al., 2009). However, particularly in the studies using text and dynamic visualizations that failed to show an effect of cueing, the cueing of relations between representations was occasionally implemented in a rather suboptimal way.

In a study by van Oostendorp and Beijersbergen (2007), the authors related a written text to highlighted parts of an animation by placing a dot in front of the corresponding paragraph. The text was given in written form, while the pace of the presentation was system-paced. It is possible that, as learners had to split their attention between text and dynamic visualizations, the learners may have missed crucial information, so that the effectiveness of cueing might have been overshadowed. Indirect support for this assumption was provided in a follow-up study by van Oostendorp et al. (2008), in which the authors used the same dynamic visualizations, but this time the presentation's pace could be determined by the learner. Results of this study showed that cued dynamic visualizations were superior to uncued ones under learner-control conditions.

In a study by Moreno (2007), labels were shown next to a dynamic visualization. These labels were highlighted when mentioned in the narration that accompanied the dynamic visualization. Results revealed no positive effect of cueing. Moreno interpreted the absence of an effect by assuming that this type of cueing may have forced learners to split their attention between the dynamic visualization and the highlighted label, so that no benefit of cueing emerged.

Mautone and Mayer (2001) investigated in a 2x2-design the presence/absence of visual cues (such as pictorial headlines and colored elements) in dynamic visualizations and the presence/absence of cues in a corresponding spoken text (i.e., adding headlines, summaries, enumerations, logical connective phrases et cetera). Results revealed only an effect for text cues, but not for visual cues in the animations. The authors gave several explanations for the lack of an effect for visual cueing. On the one hand, the complexity of the animation was considered to be rather low, making it questionable if visual cues had been necessary at all to guide the learners cognitive processing. This would also be in line with a study by Jeung et al. (1997), who found cueing to be beneficial only when the complexity of the visualizations was high. On the other hand, the cueing treatment of Mautone and Mayer (2001) may have been a rather weak one, because in the cued animation condition solely icons that acted as a kind of pictorial headline and some colored elements were used.

The same argument of a rather weak treatment might also account for the results of a study by Large et al. (1996), who could not find an effect of cueing when adding two captions to dynamic visualizations. Since in both, the cued and uncued animations, parts of the animations were labeled, the authors explained their finding by assuming that adding two captions was not necessary anymore.

Also a rather weak treatment was used in a study by Spangenberg (1973; Exp. 2). The author was comparing a narrated video to a narrated video with arrows and found no differences between conditions. The arrows were implemented not to point to relevant parts, but to show the direction of the movement. However, as the movement was already conveyed through dynamic visualizations, this function of arrows was rather redundant, so it might not be very surprising that there was no effect of cueing.

To sum up, the studies that did not show any effect of cueing in dynamic visualizations might have implemented cues in a suboptimal way, by either not adding further explanations to the animations so that cueing did not stimulate cognitive processes (de Koning et al., 2010a; Kriz & Hegarty, 2007), by using inappropriate cues such as arrows (Kriz & Hegarty, 2007; Spangenberg, 1973), by overshadowing a potential effect of cueing through a split of attention (Moreno, 2007; van Oostendorp & Beijersbergen, 2007), or by a rather weak treatment, or a lack of necessity of cues, respectively (Large et al., 1996; Mautone & Mayer, 2001; Spangenberg, 1973).

Cueing in dynamic visualizations: positive effects

In the following, studies will be described that showed a positive effect of cueing in dynamic visualizations. Studies using dynamic visualizations without text will be outlined first, before reviewing studies in which dynamic visualizations were accompanied by text.

Although de Koning et al. (2007) could find a positive effect on performance when an animation (without text) was cued as opposed to an uncued animation, these results will be neglected, as the authors could not replicate this finding for these two conditions in three further studies (cf. de Koning et al., 2010a, 2011a, 2011b).

Different presentation speeds can be regarded as a special form of cueing, since they do not add any new information, but are supposed to guide attention and make certain aspects of an animation more salient (dynamic contrast, Schnotz & Lowe, 2008). Thereby, they might make it easier for learners to select the relevant information. For instance, slow processes might become better perceivable and comprehensible in an animation by speeding the animation up. In a study by Fischer, Lowe, and Schwan (2008), different presentation speeds (fast vs. normal) were used as a cueing device for an animation of a pendulum clock. The fast presentation speed made aspects salient that in the normal speed condition were hardly perceivable. Results revealed that the different presentation speeds influenced the distribution of attention and led to differences in performance, whereby learners in the fast presentation condition outperformed the normal speed condition (see also Fischer, 2008; Exp. 2). In a follow-up study, Fischer and Schwan (2010) again compared the effectiveness of different presentation speeds of the same animation (fast vs. normal); in addition, they varied whether different parts of a pendulum clock were highlighted by blinking colors. The blinking color was supposed to guide attention via visuo-spatial contrast, that is, due to its blinking character the entity was supposed to be more easily distinguished from its surroundings (cf. Schnotz & Lowe, 2008). The different presentation speeds, on the other hand, were supposed to guide attention via dynamic contrast. The results again revealed that the fast presentation speed was beneficial for learners, whereas this was not the case for cueing different parts of the animation by means of blinking colors. Similarly, in a study by Meyer, Rasch, and Schnotz (2010), results revealed that different presentation speeds of an animation depicting a four stroke-engine affected the comprehension of macroscopic functional aspects (e.g., the timing device of the four cylinders). Even though one might argue that these results are restricted to specific characteristics of the used animations of the studies by Fischer and colleagues (2008, 2010), as well as Meyer et al. (2010), they nevertheless reveal the importance of how overemphasizing certain aspects of a motion can contribute to the understanding of a depicted topic.

Boucheix and Lowe (2010) compared two different kinds of cues, arrows and a (synchronized) spreading-color cue, with respect to their effectiveness for fostering the understanding of the mechanics of a piano. Results of their first experiment revealed that performance in the spreading-color cue condition was superior to the arrow condition and to a control condition, whereas the latter two conditions did not differ. In a second experiment, the authors compared in a 2x2-design once again type of cue (arrows vs. spreading-color cue) and whether the cues were synchronized with the time-course (synchronized vs. unsynchronized). They thereby did not only *spatially* cue the different components of the system, but also *temporally* cued the time course of the system. Note that in the condition of the unsynchronized spreading-color cue, the spreading-color cue did not cue temporally, since all colored ribbons appeared at once. This experiment confirmed the superiority of spreading-color cues to arrows, and also revealed that the synchronization of the cues to the functional time course of the system was beneficial for an understanding of the kinematics and functioning of the system. Thus, the cues might have helped the learner to organize the displayed information into a coherent mental model by guiding processing by means of synchronizing the cues according to the cause-and-effect chain.

The aforementioned studies investigated the effectiveness of cueing in dynamic visualizations without text and observed a positive effect of cueing. This is somewhat contradictory to the findings of Kriz and Hegarty (2007) as well as de Koning et al. (2010a) who could not find an effect of cueing without text and concluded that cueing might mainly stimulate perceptual processing rather than cognitive processing and might therefore not be very beneficial without further explanations. This conflicting pattern of results of cueing in dynamic visualizations without text may be explained by the functional aspects the successful cues emphasized. On the one hand, the speed of an animation was manipulated in the studies by Fischer and colleagues (2008, 2010) as well as Meyer et al. (2010). By using high speed, certain functional aspects might have become salient, which otherwise may not have really been perceivable. In the study by Boucheix and Lowe (2010), the positive effect of cueing might be traced back to the special design of this cue that emphasized the functional aspects of the system by cueing the time course of the elements, or the cause-and-effect chain, which can be considered as the most important aspect of their used instructional material. Thereby, the aforementioned cues might not only have stimulated perceptual processing by helping to select the presented information, but might also have stimulated cognitive processing (cf. de Koning et al., 2009).

In the following, research on cueing in dynamic visualizations accompanied by further explanations will be discussed. This particular method is especially interesting in the context of this thesis as the visualizations described in the current thesis were also accompanied by text.

Van Oostendorp et al. (2008) found a positive effect of cues in animations accompanied by written text. The cue within the dynamic visualization (spotlight cue) aimed at reducing the visual complexity within the dynamic visualization and at guiding attention to the relevant information in the visualization. Furthermore, there were cues that were designed to emphasize the relationship between text and visualizations. However, as already discussed above, cueing was only beneficial if the animations were self-paced, so that a split of attention between the written text and the animations could be compensated and the benefits of cueing were able to shine through.

Boucheix and Guignard (2005) implemented several cueing techniques, such as colored dots, arrows, tachometers, and verbal cues in dynamic visualizations, which in turn were accompanied by short written text consisting of one-two sentences. Results revealed that these cues aided comprehension of dynamic visualizations.

Moreover, there is evidence that cues may be helpful in learning with dynamic visualizations accompanied by spoken text.

Amadieu et al. (2011) compared an ordinary animation to an animation that was cued by zooming in on the relevant information in the visualization. Both animations were accompanied by spoken text. By zooming-in, less relevant details of the animation move out of focus. This in turn reduces the visual complexity of an animation, and, moreover, helps learners to relate information of the narration and the corresponding information of the visualization. Results showed a positive effect of cueing on learning outcomes.

In two experiments by Huk et al. (2010), cues were implemented in dynamic visualizations either by coloring elements or by adding labels, synchronized with a corresponding narration. In doing so, particularly the processes of selecting information within the visualization and making the connection with the corresponding information within the narration might be supported. For both experiments results revealed that cues increased performance in a task that required the learner to memorize important facts that were explicitly depicted in the dynamic visualizations, whereas cues had no impact on deeper understanding. The authors supposed that the latter finding might be caused by the short duration of the visual cues, and suggested that the use of prolonged as well as multiple cues may be a promising way to support more thorough comprehension.

Support for using multiple cues comes from a study by Janelle, Champenoy, Coombes, and Mousseau (2003) in which the authors showed learners a narrated video of a model performing a procedural task and highlighted relevant parts either by visual cues, verbal cues, or a combination of both. Results revealed that learners in the combined cues condition outperformed learners in the single cue conditions and in the non-cued video condition, whereas learners in the single cues condition did not outperform the pure video condition.

Jarodzka (2011; Exp. 2 & Exp. 3) used two different forms of cues across two studies: a spotlight and a dot. In both studies the cues were derived from the viewing behavior of the expert. Moreover, the narrated explanations of the expert were used as accompanying text. The cues were then synchronized with the narration. This might have not only highlighted information in the visualization, but might have also helped learners to relate the cued elements of the visualizations with the corresponding text, thereby facilitating the process of integrating these sources of information into a coherent mental model. This view is supported by the results of these two studies, since they both show that cueing guided attention to the cued parts and enhanced understanding.

De Koning et al. (2011b) investigated the influence of cues in animations. The cues either did or did not contain self-explanation prompts. A spotlight, which was supposed to deemphasize less relevant information, thereby reducing visual complexity, was used as the cueing method. The results showed that the benefits of cueing animations were only observable when learners were prompted to conduct self-explanations. In a follow-up study (de Koning et al., 2010b), cued and uncued animations that were either augmented with self-explanation prompts or with spoken instructional explanations were compared. Results indicated that cued animations were superior to uncued animations for the self-explanation conditions as well as the spoken text conditions, supporting the view that the benefits of cueing in animations might mainly become evident, if additional explanations are provided by the instructor (i.e., instructional explanations), or by the learner (i.e., self-explanations).

In conclusion, it is fair to say that studies that implemented effective cues in dynamic visualizations, which were not accompanied by text, used cues that overemphasized the motion in dynamic visualizations or that highlighted the time course of events. When accompanied with written text, cueing was mainly effective in dynamic visualizations, if these were self-paced, or if the text was short. This could possibly be accounted to the fact that a split of attention did not overshadow the potential of cues in these cases. When dynamic visualizations were accompanied by spoken text, most of the studies showed a positive effect for cueing in dynamic visualizations, indicating on a descriptive level that the potential of cueing might most likely unfold for narrated dynamic visualizations.

All in all, the mixed pattern of results for cueing in dynamic visualizations (see Table 4.4) indicates that on the one hand, cueing holds great potential for improving learning with dynamic visualizations, but that on the other hand, cueing is not a remedy in itself. Therefore, it is necessary to take a closer look at, and identify the conditions under which cueing in dynamic visualizations has proven to be beneficial. If cues were implemented in dynamic visualizations without text, their influence on comprehension might be rather subordinate, because they might primarily foster perceptual processing, but not cognitive processing (cf. de Koning et al., 2009). Nevertheless, specific designs of visual cues might also stimulate cognitive processing, for instance if they were to cue the cause-and-effect chain (e.g., Boucheix & Lowe, 2010). Moreover, there is evidence that in the studies where no effect of cueing could be observed, the potential of cues was not fully exploited (see also de Koning et al., 2009). This might especially be the case, as the movements depicted in dynamic visualizations already possess the potential to attract attention (dynamic contrast, Schnotz & Lowe, 2008), which might compete with the attention-guiding function of cues (cf. de Koning et al., 2009). Therefore, especially designing cues for dynamic visualizations might be complicated and need to be considered carefully. Based on this research overview, some guidelines might be derived for the design of successful cues in dynamic visualizations. These guidelines will address cueing techniques in learning with dynamic visualizations that are accompanied by text, since this was the case for all studies in the current thesis. To successfully cue dynamic visualizations, it might be reasonable to implement *multiple cues* that on the one hand support the processes of selection, organization, and integration (cf. de Koning et al., 2009; Mayer & Moreno, 2003), and counteract the assumed visual complexity of dynamic visualizations on the other hand.

First, cues should facilitate the selection of relevant information. Doing so, they might overemphasize dynamic features, so that these aspects become more salient (e.g., Fischer, 2008), and, furthermore, they might overshadow the dynamic contrast of less relevant and distracting movements, for instance by means of a spotlight (e.g., de Koning et al., 2007, 2010a, 2010b, 2011a, 2011b). Thereby, especially the latter treatment is supposed to reduce the visual complexity of dynamic visualizations, since less relevant elements decrease distraction. Second, cues should help in organizing the information (depicted by the visualizations) into a coherent pictorial mental model. This might be realized by the use of cues that correspond to the temporal order of a system, or to the cause-and-effect chain, respectively (e.g., Boucheix & Lowe, 2010). Since in this case, the order of processing elements within a visualization can also be guided, so that learners know when to attend to which location, the problems arising from an intra-representational split-attention effect, and in turn, the visual complexity of dynamic visualizations, might become extenuated. Third, cues should aid learners in making the connection between certain elements of the text and their corresponding elements within a visualization, thereby supporting the process of integrating information from text and visualizations into a coherent mental model (cf. de Koning et al., 2009). With spoken text, for instance, this could be realized by adding elements to dynamic visualizations step by step, when each is first mentioned in the text (e.g., Huk et al., 2010). Once again, such a cueing method might additionally reduce the problems arising from the visual complexity of dynamic visualizations, particularly problems associated with an intra-representational split of visual attention, since the visualizations gradually build up. Another way to help learners to integrate information from text and visualization might be to synchronize the cued part of the visualization with the verbal explanation (e.g., Amadieu et al., 2011; Huk et al., 2010; Jarodzka, 2011; Exp. 2 & Exp. 3).

In compliance with these suggestions, Study 3 of this thesis made use of multiple cues in dynamic visualizations. These cues aimed at supporting the processes of selection, organization, and integration, and also aimed at counteracting problems associated with the visual complexity of dynamic visualizations. Note that arrows as a type of cueing device were omitted, as their benefits are highly questionable for cueing dynamic visualizations (e.g., Boucheix & Lowe, 2010). A detailed description of the cues used is given in the method section of Study 3 of the current thesis.

In the reminder of this chapter conclusions about the conducted literature overview of cueing in static visualizations and cueing in dynamic visualizations will be drawn.

The role of cueing in learning with dynamic as opposed to static visualizations

At first glance, the effectiveness of cueing in static visualizations seems to be overall more successful than in dynamic visualizations. However, the research on how to construct successful cues for dynamic visualizations is somewhat in its early stages (cf. de Koning et al., 2009). This might be traced back to the fact that the construction of helpful cues might be easier for static visualizations than for dynamic visualizations: In dynamic visualizations the attention guiding function of cueing has to compete with the dynamic contrast of dynamic visualizations, which in turn is supposed to attract attention.

However, when carefully designing cues for dynamic visualizations, cues might counteract the factors that are supposed to constitute the high degree of visual complexity in dynamic visualizations and therefore unfold the potential of dynamic visualizations. For instance, a spotlight can de-emphasize less relevant distracting elements, or gradually building up the visualization can reduce the problems arising from intra-representational split-attention. Because especially dynamic visualizations compared to static visualizations may possess a comparatively high degree of visual complexity, cueing might be even more helpful in learning with dynamic than with static visualizations.

Two studies are known to have attempted to investigate the influence of cues on learning with dynamic and static visualizations: Spangenberg (1973) as well as Boucheix and Guignard

(2005). However, these studies have flaws that have already been partly discussed above. Spangenberg (1973; Exp. 2) used arrows as a cueing device to indicate the direction in which the movement would lead. The results of this study revealed a main effect of type of visualizations in favor of dynamic visualizations, but no main effect for cueing and no interaction. However, as in dynamic visualizations the movement is already depicted, one would not expect cues to be helpful there. If anything, these cues should be mainly helpful for static visualizations, where they help to convey information concerning direct movement. In the study by Boucheix and Guignard (2005), results revealed a main effect for type of visualization in favor of dynamic visualizations, and a main effect of cueing, but no interaction. The authors implemented multiple cues, among those also tachometers and verbal cues. A tachometer explicitly depicts information about velocity and changes in thereof. However, the tachometer might be especially helpful for learners in the static visualization conditions, who heavily rely on dynamic information from other sources, such as text (Kühl, Scheiter, & Gerjets, 2010). Moreover, learners in the dynamic visualizations condition can directly read-off dynamic information and, hence, may have even been unnecessarily distracted by the tachometer, which, taken together, may have overshadowed a potential interaction. In addition, the implemented verbal cues consisted of prompts on how to process the visualizations. This is problematic for two reasons: First, such prompts are not regarded as cueing anymore, but may rather be seen as an instructional learning strategy. Second, while learners in the dynamic visualizations conditions were prompted to compare the speeds of two gears by perceiving their speeds, learners in the static visualizations conditions were prompted to compare the speeds of two gears by inferring their speeds. Accordingly, it is unclear if the added value of the implemented cueing methods in this study can be traced back to what can be considered literal cueing, or if it might be traced back to prompting learners to engage in certain cognitive processes, which also differed from each other (i.e., prompt to perceive and compare speeds with dynamic visualizations vs. prompts to infer and compare speed with static visualizations).

To conclude, research predominantly shows that cueing static visualizations is beneficial for learning. With regard to cueing dynamic visualizations, results are at first glance inconclusive, indicating that cueing dynamic visualizations is not a remedy in itself. Rather, it might be the case that cues have to be implemented in a way that they are able to counteract the problems associated in learning with dynamic visualizations, particularly their visual complexity, and, furthermore, support the processes of selection, organization, and integration. Here, such carefully designed cues may be a promising way to optimize learning with dynamic visualizations. Moreover, as such cues aim to cope with the processing demands associated with a high degree of visual complexity – which in turn is assumed to be more harmful in dynamic than in static visualizations – these cues may be even more beneficial for dynamic than for static visualizations. Finally, with respect to static visualizations, seeing as static-simultaneous visualizations may suffer more from a higher degree of visual complexity than static-sequential ones, the benefits of cueing might be more pronounced in static-simultaneous visualizations. These assumptions will be investigated in Study 3[22].

[22] Note that due to the way cueing was implemented in the abovementioned studies that examined the influence of cueing in dynamic and static visualization (Boucheix & Guignard, 2005; Spangenberg, 1973), these studies are inconclusive with respect to the research questions of Study 3 for the current thesis. Moreover, Study 3 is different from the other two studies with regard to other design factors and research questions, respectively (e.g., the impact of cueing on different presentation formats of static visualizations).

4.3 Conclusions

To sum up, there are two problems that might specifically account for learning with dynamic visualizations as compared to static visualizations: An inter-representational split-attention effect and the visual complexity of dynamic visualizations themselves.

To overcome inter-representational spilt-attention, and to foster learning, it might be reasonable to use spoken text. Using spoken text instead of written text should – in terms of CTML – mainly have a positive impact on the pictorial, as well as the integrated mental model, and hence mainly affect performance in pictorial tasks and transfer tasks. Using spoken text should also reduce working memory demands, thus reducing ECL. Furthermore, the benefits of dynamic visualizations as compared to static visualizations might be even more pronounced for spoken than for written text, as the potential of dynamic visualizations to convey a deeper understanding of the content might particularly unfold when problems arising from an inter-representational spilt-attention effect are solved. As the superiority of dynamic over static visualizations is solely assumed to become evident for transfer tasks, the moderating role of the text modality should also mainly affect transfer tasks. These research questions will be addressed in Study 2.

However, using spoken text does not solve the problem of a high degree of visual complexity within dynamic visualizations. To cope with this, it has been suggested to optimize learning with dynamic visualizations by means of cueing. However, recent research shows that cueing is not a remedy in itself. It may actually be necessary to carefully design cues in order to counteract the assumed high degree of visual complexity in dynamic visualizations. Such carefully designed cues might specifically help learners to build a better pictorial as well as integrated mental model, and should therefore mainly affect pictorial and transfer tasks. Cueing should also reduce unnecessary processing demands (i.e., decrease in ECL), and may stimulate learners to engage in more valuable processing activities (i.e., increase in GCL). Moreover, since particularly dynamic visualizations might suffer from a high degree of visual complexity, the impact of cueing might be more pronounced for learning from dynamic than from static visualizations. Thereby, the potential of dynamic visualizations for fostering a deeper understanding might specifically come true, and should mainly affect transfer tasks, as they are assumed to be an indicator for a deeper understanding. With regard to different presentation formats of static visualizations, it might be argued that the benefits of cueing may be more pronounced for static-simultaneous as opposed to static-sequential visualizations as static-simultaneous visualizations are supposed to be more visually complex. These assumptions will be examined in Study 3.

5 Study 2: The Influence of Text Modality on Learning with Static and Dynamic Visualizations[23]

5.1 Research Question and Hypotheses of Study 2

Learning with text and visualizations might be enhanced when problems arising from inter-representational split-attention are solved. In Study 1 of this thesis, however, inter-representational split-attention may have been existent, since the text in Study 1 was presented in written form, thereby requiring learners to shift their attention between reading the text and processing the picture. Written text had to be used in Study 1, because think-aloud protocols were assessed, while learners were dealing with the multimedia instruction. As explicated in Chapter 4.1, to overcome problems arising from inter-representational split-attention, it is recommended to use spoken instead of written text (cf. modality effect; e.g., Ginns, 2005; Low & Sweller, 2005; Mayer, 2009; Sweller et al., 1998).

Accordingly, a main effect for learning outcomes in favor of spoken text as compared to written text was expected. More precisely, this was assumed to be the case for pictorial recall tasks as well as transfer tasks, but not necessarily for verbal factual knowledge tasks, since spoken text should mainly have an impact on the pictorial as well as integrated mental model. Correspondingly, spoken text as opposed to written text should lead to a decrease of ECL, thereby leaving more resources available for GCL.

Even though in Study 1 no differences concerning learning outcomes between dynamic and static visualizations were observable, in the current Study 2 it was nonetheless assumed that for this modified and enhanced instructional material, dynamic visualizations would be superior to static visualizations with respect to a deeper understanding of the underlying dynamics. This was expected to be the case for at least three reasons.

First, the dynamic visualizations were improved in that dynamic interrelations became more evident (e.g., the interplay of changes in the frequency of the undulatory movement, its impact on changes in the magnitude of the reaction force, and the interrelatedness with swimming speed).

Second, learners had to rely more on the visualizations, since, in contrast to Study 1, visuo-spatial aspects were mainly depicted in the visualizations, but not in the text anymore.

[23] This chapter is based on: Kühl, T., Scheiter, K., Gerjets, P., & Edelmann, J. (2011). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior, 27*, 29-35.

Thereby, because learners had to rely more on the visualizations, differences between the types of visualizations could become more evident. Moreover, even though dynamic features were still described in the text of Study 2, these descriptions were less extensive compared to Study 1. On the one hand, this reduction of dynamic features in the text, which are already depicted in dynamic visualizations, can be regarded as a reduction of redundancy, and, in turn may be beneficial for learners receiving dynamic visualizations (cf. Schmidt-Weigand & Scheiter, 2011). On the other hand, for learners receiving static visualizations, the reduction of dynamic features of the text might impose higher demands in constructing an adequate mental model of the underlying dynamics. In sum, this change of text content might favor learning with dynamic visualizations.

Third, learner control over the learning environment was eliminated, so that the effectiveness of learning with dynamic and static visualizations could take place under more controlled conditions. However, learner control might rather help learners receiving static visualizations, since drawbacks in learning with static visualizations might be compensated through a more extensive use of the learning environment, as indicated by the results of Study 1. Thus, also eliminating learner control might have favored learning with dynamic visualizations.

To conclude, with respect to type of visualization (static vs. dynamic), it was expected that a main effect for learning outcomes in favor of dynamic visualizations would be observed, particularly for transfer tasks, as differences between these types of visualizations are supposed to manly affect tasks asking for a deeper understanding (cf. Bétrancourt & Tversky, 2000). In connection, dynamic visualizations were assumed to decrease ECL, since learners receiving dynamic visualizations do not need to engage in resource-demanding processes, like spatial and temporal inferences. Thereby, dynamic visualizations may leave more working memory resources available for engaging in helpful learning activities, which are associated with an increase in GCL.

Most importantly, an interaction between type of visualization and text modality was expected for learning outcomes and cognitive load: The advantage of learning with dynamic as opposed to static visualizations might be even more accentuated when using spoken instead of written text, since the potential of dynamic visualizations might particularly unfold with spoken text. This should be the case, because problems associated with a high degree of visual complexity in dynamic visualizations might be less severe when learners do not have to split their attention between text and visualizations (as for spoken text), so that the information depicted in dynamic visualizations may be processed appropriately. The assumed moderation should mainly affect transfer tasks – since also differences between dynamic and static visualizations are expected to become solely evident for transfer tasks – and should be mirrored by the respective patterns of ECL and GCL.

Finally, it was examined, if spatial abilities would moderate learning with dynamic and static visualizations. It was supposed that especially learners with weaker spatial abilities should profit from dynamic visualizations, whereas for learners with stronger spatial abilities the benefit of dynamic visualizations was expected to be less pronounced.

5.2 Method

5.2.1 Participants and Design

Eighty students with various educational backgrounds from the University of Tuebingen, Germany, participated for either course credit or payment in the study. Students had to be native speakers and not to study physics. Sixty-three female and 17 male participants (*M* = 24.39 years, *SD* = 2.97) were randomly assigned to one of four conditions, which resulted from a 2x2-design with text modality (spoken vs. written) and visualization type (static vs. dynamic) as independent variables.

5.2.2 Instructional Materials

The instructional material was a revised version of the used material in Study 1. Major changes concerned the following points: First, the redundancy of text and visualizations was reduced, in that the text itself contained only little information concerning visuo-spatial relations; these aspects were depicted in the corresponding visualizations, as has been shown to be beneficial (Schmidt-Weigand & Scheiter, 2011). Moreover, the description of dynamic features in the text was reduced to avoid a redundancy between dynamic visualizations and text. Second, particularly the dynamic visualizations were improved by using their potential to depict dynamic changes directly. At this, also background had to be added to the visualizations. Third, the instructional material was now presented system-paced and not self-paced anymore[24]. Finally, mainly on the basis of the advice from two science educators who used similar material in high school, as well as on the basis of the verbal protocols obtained from Study 1, refinements of the content were made for Study 2 in that some subtopics were explained in a more comprehensible way (e.g., the

[24] This was done, because on the one hand, learners in Study 1 played less frequently the dynamic visualizations than their counterparts in the static visualization condition. However, to be better able to control for the latter one, it was decided to investigate the effectiveness of learning with dynamic and static visualizations under more controlled conditions by keeping learning time constant by means of system-pacing.

interplay of velocity of the propelling element and its impact on the resulting force) and some topics were eliminated to not overcomplicate matters (e.g., calculating the reaction force by taking the mass of the displaced water into account).

The computerized instructional material dealt with the physical principles underlying fish locomotion and consisted of seven sections that dealt with: 1) Swim styles, 2) Propelling element, 3) Undulatory motion, 4) Actio and Reactio, 5) Magnitude of the reaction force, 6) Decomposing the reaction force, 7) Interaction of forces of various propelling elements. At this, the instructional material was focusing on various aspects of the locomotion by explaining the interplay of the trajectory and velocity of different body parts, the sizes of the associated resulting forces (symbolized by lengths of arrows), and the related swimming speed. For instance, two different swimming speeds were chosen to depict the relation among the frequency of the movement of the body parts, the associated changes in the sizes of the resulting forces, and the related swimming speed. Compared to the natural velocity of the movement of the body parts, the speed of the dynamic visualization was slowed down, as otherwise learners would not have been able to perceive the movements adequately. The decision for the chosen speeds was based on a consensus reached with the local domain experts of the larger research project (i.e., two marine biologists).

The instructional material was presented system-paced with each of the seven sections lasting 45 seconds, corresponding to the length of the spoken text for each section. The text contained 520 words and was presented – depending on the experimental condition – in either spoken or in written form. The written text was presented on the left half of the screen, while for the spoken conditions the left half of the screen remained blank. Spoken text was presented on ear phones. The text was spoken by a skilled female voice. Speech rate was moderate (approx. 3.5 syllables per second, respectively 1.8 words per second) and also the complexity of the whole text in terms of readability was moderate (Flesch-Index = 30).

In the condition with dynamic visualizations, an animation presented the undulatory movement of a fish in a repetitive fashion, meaning that the movement of the swimming fish was looped. The animation depicted the same fish across the different instructional sections. In contrast to the dynamic visualizations used in Study 1, the dynamic visualizations used in Study 2 were improved by depicting additional dynamic features. More precisely, the interplay of the changes in the frequency of the undulatory movement that are associated with changes in the magnitude of the reaction force (symbolized by the length of arrows), as well as their interrelatedness with swimming speed were depicted in the dynamic visualizations that were used in Study 2. Particularly for representing the swimming speed, landscape background was added to provide learners a reference point that allows realizing the changes in the fish's swimming speed (cf. Figure 5.1).

The static visualization condition consisted of nine key frames per section that were extracted from the corresponding animations. The key frames were displayed sequentially, whereby the nine static key frames represented two loops of an undulatory movement (see Figure 5.1). As each section lasted 45 seconds, each key frame of a section remained visible on the screen for five seconds.

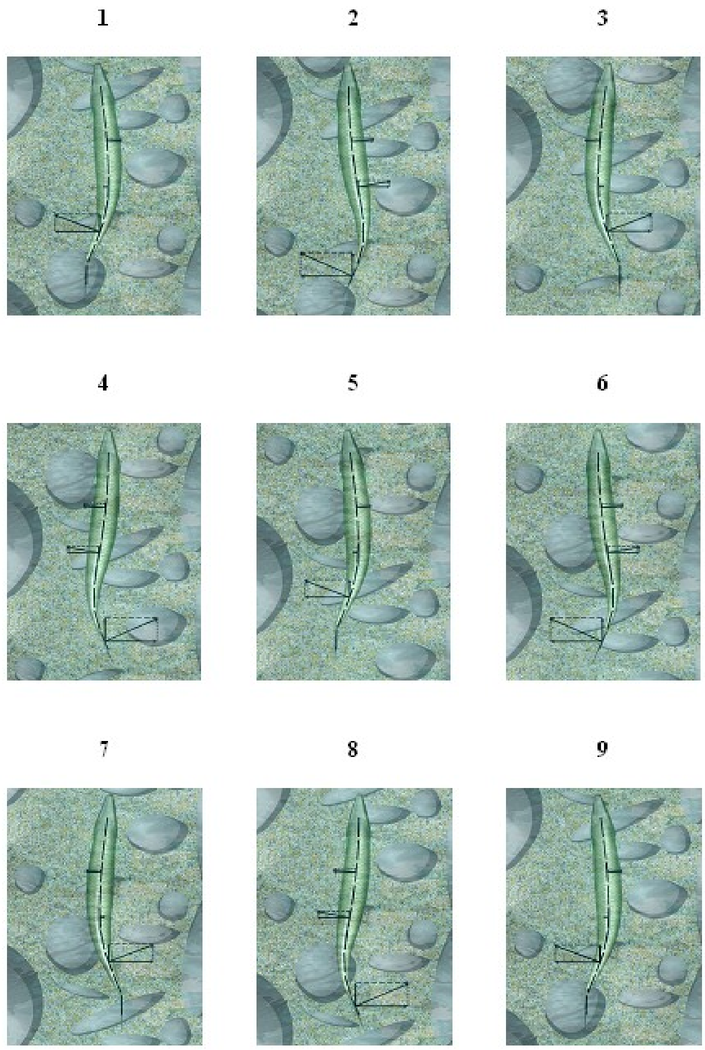


Figure 5.1. Sequence of nine static key frames in the static visualization conditions.

5.2.3 Measures

A questionnaire with respect to the attitudes towards biology and physics, as well as a prerequisite knowledge test served as control variables and a spatial ability test served as a moderator variable (and also as a control variable). As dependent variables served on the one hand three items asking about the experienced cognitive load during the learning phase, and furthermore several knowledge tests to measure learning outcomes.

Attitudes towards biology and physics. The questionnaire concerning the attitude towards biology and physics of Study 1 was used with two modifications: First, instead of seven items for each scale, only five items of each scale were used, namely the five items with the best loadings on the respective factors (see Appendix B), and secondly, instead of a 5-point Likert scale, a 4-point Likert scale was used ranging from 1 ("I strongly disagree") to 4 ("I strongly agree").

Prerequisite knowledge test. Compared to Study 1, the prerequisite knowledge test was extended with three more questions to a total of eight multiple-choice questions. The eight questions were asking for the second and third Newton axioms, the physical definition of forces, the characteristics of a harmonic oscillation, knowledge about velocity and acceleration, knowledge about the undulatory swimming style, and knowledge about the interplay of forces by analyzing a parallelogram of forces (see below for a sample item). A person's knowledge about these basic definitions and principles was considered a beneficial prerequisite for more easily achieving an understanding of the topics explained in the current study. Hence, it was not the aim at testing for a deeper understanding of physical concepts or principles that were unlikely to be present in the used student sample. The eight multiple-choice questions consisted of four to six alternatives to choose from and for each question there were one to three correct answers. For each correct answer, learners were assigned one point and for each wrong answer one point was subtracted. Within a question, however, learners could at worst receive zero points. The maximum score was 12 points.

Example of a question from the prerequisite knowledge test

Which of the following correctly describes Newton's third law of motion?

a) Gravity causes objects to fall towards the earth's centre.

b) When body A exerts force on to body B, a force of the same magnitude, acting in the opposite direction, results from B onto A.

c) Every action results in a smaller reaction.

d) A body remains at equilibrium due to the action and reaction principle.

Spatial ability test. To control for individual differences in spatial abilities and to examining their potential moderating role, the mental rotation test (MRT) was administered (Vandenberg & Kuse, 1978). The MRT was used, as especially learning with the static visualizations used in this study required the ability to mentally rotate and manipulate visuo-spatial objects (e.g., to imagine the movement of the caudal fin). The MRT consists of 20 items, whereby each item comprises a complex three-dimensional block figure and four alternative figures as multiple-choice answer options. For each item, the participant has to choose, which two of the four alternative figures are identical to the target when (mentally) rotated. There is a time limit of six minutes for working on the MRT. For each correctly identified figure one point was given and for each wrong identified figure one point was subtracted, resulting in a maximum of 40 points and a minimum of -40 points.

Cognitive load measures. Because the perceived difficulty item ("How difficult was it for you to understand the contents?"), which was assessed in Study 1 seemed to be suited to measure ECL, it was also assessed in the current Study. This item will be named ECL_1-item in the following. Furthermore, since Cierniak, Scheiter, and Gerjets (2009) could identify in their study two items that successfully measured ECL and GCL, respectively, also these two items were assessed. The item supposed to measure ECL was: "How difficult was it for you to learn with the given material?"(original German version "Wie schwer ist es dir gefallen, mit dem gegebenen Material zu lernen?"), and will be named ECL_2-item in the following. The item supposed to measure GCL was: "How much did you concentrate during learning?" (original German version: "Wie sehr hast du dich während der Lernphase konzentriert?"). Each of these three items had to be rated on a nine-point Likert scale.

Knowledge tests. In comparison to Study 1, the content of the instructional material for Study 2 was partly changed and reduced. Accordingly, also the knowledge test was changed in that some items were eliminated, some new items were added, and some other items were changed to fit to the new instructional material. Thereby, learning outcomes were measured by means of six multiple-choice questions assessing verbal factual knowledge, five pictorial recall tasks, and eleven transfer tasks (see below for sample items of each test). A maximum of 11 points could be achieved for the verbal factual knowledge test, a maximum of 9 points could be achieved for the pictorial recall test, and a maximum of 26 points could be achieved for the transfer test. The verbal factual knowledge questions were posed in written form and all correct answers to these tasks had been explicitly conveyed by the multimedia instruction. The pictorial recall tasks were posed in pictorial form and the correct answers to the tasks were essentially conveyed by the visualizations, but not by the text. The transfer tasks were posed in written as well as in pictorial form. To solve the transfer tasks, learners had to apply their acquired knowledge to new scenarios.

The open questions were corrected by two independent raters. Cases of disagreement (3.34%) were resolved by discussion.

Example of a question from the verbal factual knowledge test

Which of the following is/are true?

a) The reaction force can be broken down into a propelling force and a lateral force.

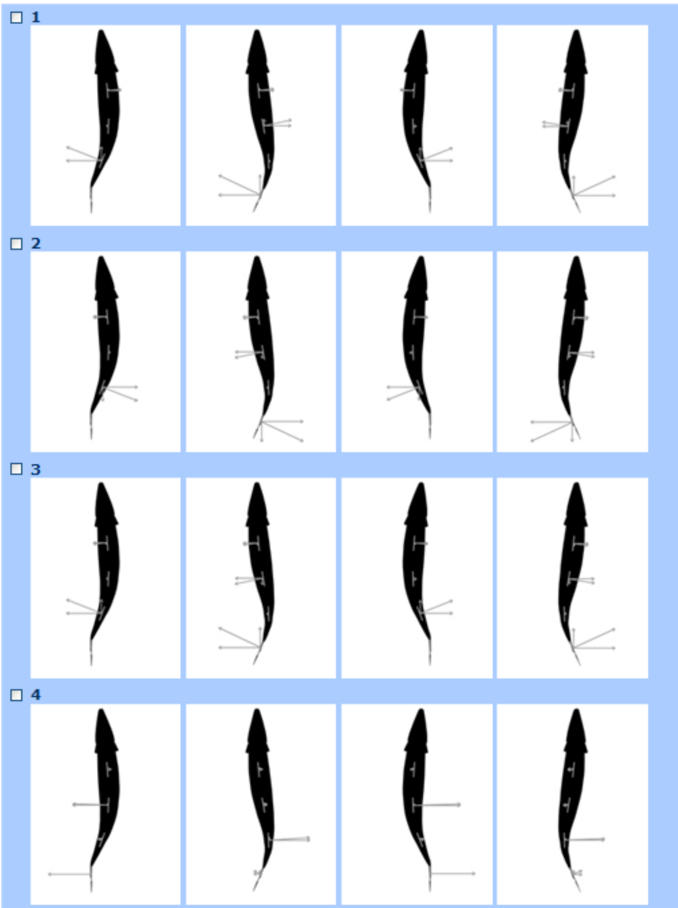
b) Lateral force and reaction force are perpendicular to each other.

c) Propelling force and lateral force are perpendicular to each other.

d) Reaction force and propelling force are perpendicular to each other.

Example of a question from the pictorial recall test

Which of the following sequences correctly illustrates the forces emerging during undulation?



Example of a question from the transfer test

Some undulating species of fish move their head back and forth in order to swim forwards. Why is this? Write down any feasible reasons you can think of!

5.2.4 Procedure

Each participant was tested individually during a session of approximately 90 minutes. First, the questionnaire concerning attitudes towards biology and physics, then the prerequisite knowledge test and then the mental rotation test were administered. Thereafter, the subjects were presented with the learning materials and were subsequently asked to rate their cognitive load experienced during learning. Finally, the subjects took the knowledge tests.

5.3 Results

First, the questionnaire concerning attitudes towards biology and physics was analyzed by means of a factor analysis, to validate the assumed independence of the two factors. Then students' attitudes towards biology and physics, their prerequisite knowledge as well as their spatial ability scores were analyzed by means of ANOVAs to test if the experimental conditions could be regarded as equal with respect to these influencing variables. Then the abovementioned hypotheses were tested. The procedure to test if spatial abilities moderated learning outcomes in the two visualization conditions was as follows: The scores of the spatial ability test for participants of the two visualization conditions were z-standardized and used as a continuous factor in the respective ANCOVAs. Partial eta-squared ($\eta^2 p$) is reported as measures of effect size.

5.3.1 Comparability of Experimental Conditions with Respect to Attitude Towards Biology and Physics, Prerequisite Knowledge, and Spatial Abilities

Since the attitude questionnaire slightly changed from Study 1 to Study 2, once again a factor analysis was conducted for the questionnaire used in Study 2. Because two different scales were assumed, namely a biology scale and a physics scale, two factors were extracted by principle component analysis, and rotated by varimax rotation. Note that negatively formulated items were recoded. Loading of items on factors are depicted in Appendix B. Items are grouped by factors and by size of loading to facilitate interpretation. As can be seen in Appendix B, the items loaded on the one hand well on their assumed factors with all loading > .60 (cf. Bortz, 2005), and, moreover, poorly on the other factor (all loadings < .30).

The internal consistency for the shortened version of the biology scale and physics scale were still excellent, with $\alpha = .90$ for the biology scale, and $\alpha = .91$ for the physics scale. Because of the support of the two scales by means of the factor analysis as well as their high internal consistency, the items of each scale were subsumed to one biology score and one physics score, respectively. Means and standard deviations are reported in Table 5.1. With respect to the biology scale, there were no statistically significant differences for type of visualization ($F(1, 76) = 2.41$, $MSE = 11.00$, $p = .13$, $\eta^2p = .03$), text modality ($F(1, 76) = 1.82$, $MSE = 11.00$, $p = .18$, $\eta^2p = .02$), and no interaction ($F < 1$, *ns*). Also for the physics scale, there were no significant differences for type of visualization, text modality, and no interaction (all $Fs < 1$, *ns*). Concerning prerequisite knowledge, there were no statistically significant differences for type of visualization or text modality (both $Fs < 1$, *ns*). However, there was a significant interaction with regard to prerequisite knowledge ($F(1, 76) = 3.93$, $MSE = 270.42$, $p = .05$, $\eta^2p = .05$): For written text, there were no differences between dynamic and static visualizations ($F < 1$, *ns*), whereas for spoken text, there was a marginal significant effect ($F(1, 76) = 3.40$, $MSE = 270.42$, $p = .07$, $\eta^2p = .04$), with learners in the static visualization condition possessing higher prerequisite knowledge than learner in the dynamic visualization condition. For dynamic visualizations, there was no difference between the spoken and the written text condition ($F(1, 76) = 1.26$, $MSE = 270.42$, $p = .27$, $\eta^2p = .02$), whereas for static visualizations, there was a marginal effect ($F(1, 76) = 2.83$, $MSE = 270.42$, $p = .097$, $\eta^2p = .04$), with learners in the spoken text condition possessing higher prerequisite knowledge than learners in the written text condition. With respect to spatial abilities, there was no main effect for text modality ($F < 1$, *ns*), and no interaction ($F(1, 76) = 1.36$, $MSE = 78.06$, $p = .25$, $\eta^2p = .02$). However, there was a marginal significant effect for type of visualizations showing that learners in the dynamic visualization condition tended to possess better visuo-spatial abilities ($F(1, 76) = 3.70$, $MSE = 78.06$, $p = .06$, $\eta^2p = .05$). As prerequisite knowledge was not equal among conditions, it was used as a covariate in all analyses reported in this Study. Since spatial ability is inserted in the analyses as a continuous factor to test its moderating role anyway, differences among the visualization conditions are controlled for.

Table 5.1

Means (and SD) as a Function of Type of Visualization and Modality

	Dynamic visualizations		Static visualizations	
	Spoken (n = 20)	Written (n = 20)	Spoken (n = 20)	Written (n = 20)
Attitude towards biology-scale (5-20)	16.95 (3.09)	18.25 (3.04)	16.10 (4.09)	16.80 (2.91)
Attitude towards physics-scale (5-20)	11.95 (3.56)	12.30 (4.04)	13.10 (3.34)	12.45 (3.82)
Prerequisite knowledge (% correct)	50.83 (16.64)	56.67 (16.58)	60.42 (15.50)	51.67 (17.01)
Spatial abilities (-40 – 40)	18.50 (5.85)	20.30 (10.26)	17.00 (7.53)	14.20 (10.77)
Cognitive load (1-9)*				
ECL_1	4.82 (.42)	4.36 (.43)	5.28 (.42)	5.34 (.43)
ECL_2	4.64 (.42)	4.50 (.42)	5.37 (.42)	5.39 (.43)
GCL	7.27 (.29)	8.08 (.29)	6.86 (.29)	7.65 (.30)
Learning outcomes (% correct)*				
Verbal factual knowledge	54.59 (3.80)	58.23 (3.86)	55.41 (3.81)	52.95 (3.88)
Pictorial recall	44.83 (3.88)	40.22 (3.94)	46.25 (3.89)	35.04 (3.97)
Transfer	49.57 (2.39)	45.56 (2.43)	43.35 (2.40)	36.50 (2.44)

* Learning outcomes, as well as cognitive load ratings, are adjusted by taking into account prerequisite knowledge as a covariate, and spatial abilities as a continuous variable; values in parentheses refer to standard errors for these dependent measures.

5.3.2 Effects of Text Modality, Visualization Format, and Spatial Abilities on Learning Outcomes

Two-factorial ANCOVAs with text modality and type of visualization as between subject factors, with using prerequisite knowledge as a covariate were conducted for the dependent variables verbal factual knowledge, pictorial recall, and transfer knowledge. Furthermore, to test if spatial abilities moderated learning with the two types of visualizations, the scores of the MRT were z-standardized and their interaction with type of visualization was inserted into the respective ANCOVAs.

One-factorial ANCOVAs revealed no effects of type of visualization for verbal factual knowledge, or pictorial recall (both $Fs < 1$, *ns*), but they did for transfer ($F(1, 73) = 9.92$, MSE =

111.06, $p < .01$, $\eta^2 p = .12$), indicating that learners presented with dynamic visualizations performed better in the transfer tasks.

Concerning modality, the one-factorial ANCOVAs for the three learning outcome measures revealed no differences for verbal factual knowledge ($F < 1$, *ns*), but a significant effect for pictorial recall ($F(1, 73) = 4.17$, $MSE = 291.55$, $p = .045$, $\eta^2 p = .05$), as well as for the transfer test ($F(1, 73) = 5.20$, $MSE = 111.06$, $p = .03$, $\eta^2 p = .07$), indicating that learners presented with spoken rather than written text performed better in the latter two tests.

With respect to the interaction of type of visualization and modality, neither the ANCOVA for verbal factual knowledge, nor for pictorial recall, nor for transfer revealed significant results (all $Fs < 1$, *ns*).

Concerning spatial abilities, the one-factorial ANCOVAs for the three learning outcome measures revealed no differences for verbal factual knowledge ($F(1, 73) = 2.24$, $MSE = 285.60$, $p = .14$, $\eta^2 p = .03$), but a significant effect for pictorial recall ($F(1, 73) = 13.88$, $MSE = 291.55$, $p < .001$, $\eta^2 p = .16$), as well as for the transfer test ($F(1, 73) = 7.77$, $MSE = 111.06$, $p < .01$, $\eta^2 p = .10$). The effects indicated that higher visuo-spatial abilities were associated with better performance in pictorial recall ($N = 80$; $r = .44$; $p < .001$) as well as transfer ($N = 80$; $r = .38$; $p = .001$). However, other than expected, there was no interaction of spatial abilities and type of visualizations for any of the learning outcome measures, neither for verbal factual knowledge ($F(1, 73) = 2.22$, $MSE = 285.60$, $p = .13$, $\eta^2 p = .03$), nor for pictorial recall ($F < 1$), nor for transfer ($F(1, 73) = 1.05$, $MSE = 111.06$, $p = .30$, $\eta^2 p = .01$), indicating that spatial abilities did not moderate learning with the different types of visualizations.

With regard to prerequisite knowledge, the one-factorial ANCOVAs for the three learning outcome measures revealed no effects for the transfer test ($F(1, 73) = 1.64$, $MSE = 111.06$, $p = .21$, $\eta^2 p = .02$), but a marginally significant effect for verbal factual knowledge ($F(1, 73) = 3.90$, $MSE = 285.60$, $p = .052$, $\eta^2 p = .05$), as well as a significant effect for pictorial recall ($F(1, 73) = 7.86$, $MSE = 291.55$, $p < .01$, $\eta^2 p = .10$). The effects indicated that a higher prerequisite knowledge was associated with better performance in verbal factual knowledge ($N = 80$; $r = .26$; $p = .02$) as well as pictorial recall ($N = 80$; $r = .36$; $p = .001$).

5.3.3 Effects of Text Modality, Visualization Format, and Spatial Abilities on Cognitive Load

Two-factorial ANCOVAs with text modality and type of visualization as between subject factors, with using prerequisite knowledge as a covariate were conducted for the dependent variables ECL_1, ECL_2, and GCL. Furthermore, to test if spatial abilities moderated experienced cognitive load when learning with the two types of visualizations, the scores of the MRT were z-standardized and their interaction with type of visualization was inserted into the respective ANCOVAs.

For the ECL_1-item, which was also used in Study 1, a 2x2-ANCOVA showed a marginal main effect for type of visualization ($F(1, 73) = 2.87$, $MSE = 3.43$, $p = .09$, $\eta^2 p = .04$) with learners presented with static visualizations indicating higher ECL ratings than learners presented with dynamic visualizations. However, there was neither a main effect for modality nor an interaction between modality and type of visualizations, nor an interaction between spatial abilities and type of visualizations (all $Fs < 1$, *ns*). Moreover, there was also no effect for spatial abilities, or for prerequisite knowledge (both $Fs < 1$, *ns*).

For the ECL_2-item, which was adopted from the Study by Cierniak et al. (2009), a 2x2-ANCOVA showed a marginal main effect for type of visualization ($F(1, 73) = 3.65$, $MSE = 3.39$, $p = .06$, $\eta^2 p = .05$) with learners presented with static visualizations giving higher ECL ratings than learners presented with dynamic visualizations. However, there was neither a main effect for modality nor an interaction between modality and type of visualizations, nor an interaction between spatial abilities and type of visualizations (all $Fs < 1$, *ns*). Moreover, there was also no effect for spatial abilities ($F < 1$, *ns*), or for prerequisite knowledge ($F(1, 73) = 2.62$, $MSE = 3.39$, $p = .11$, $\eta^2 p = .04$).

A 2x2-ANCOVA for the GCL-item revealed a main effect of modality ($F(1, 73) = 7.72$, $MSE = 1.62$, $p < .01$, $\eta^2 p = .10$), with learners in the written text condition giving higher GCL ratings than learners in the spoken text condition. There was no main effect for type of visualization ($F(1, 73) = 2.10$, $MSE = 1.62$, $p = .15$, $\eta^2 p = .03$) and no interaction between modality and type of visualizations ($F < 1$, *ns*). There was also no effect for spatial abilities, or for prerequisite knowledge (both $Fs < 1$, *ns*). Moreover, the interaction between spatial abilities and type of visualizations failed to reach statistical significance ($F(1, 73) = 2.80$, $MSE = 1.62$, $p = .10$, $\eta^2 p = .04$).

While there was a strong positive relationship between the ECL_1-item and the ECL_2-item ($r = .56$, $p < .001$), there was no relationship between the ECL_1-item and the GCL-item ($r = .04$, $p = .72$), or the ECL_2-item and the GCL-item ($r = -.13$, $p = .26$). Higher ratings on the ECL_1-item were associated with lower performance on transfer tasks, while higher ratings on the ECL_2-item were associated with lower learning outcomes for verbal factual knowledge as well as transfer. Moreover, higher GCL ratings were associated with higher verbal factual knowledge (see Table 5.2).

Table 5.2

Correlations Among Cognitive Load Measures and Knowledge Tests

n = 80	Verbal factual knowledge	Pictorial recall	Transfer knowledge
ECL_1	$r = -.12$	$r = -.05$	$r = -.22^*$
ECL_2	$r = -.24^*$	$r = -.12$	$r = -.23^*$
GCL	$r = .24^*$	$r = .11$	$r = -.01$

Note: * $p < .05$

5.4 Summary and Discussion

In the current study, it was investigated whether multimedia learning in general, and learning from dynamic visualizations in particular, could be optimized by using spoken text for presenting instructional explanations.

First, it was expected that learning with dynamic as opposed to static visualizations would result in better learning outcomes, particularly for transfer tasks, a decrease in ECL, and an increase in GCL. Partial support in line with these assumptions was obtained in the study. That is, dynamic visualizations proved to be superior to static visualizations for transfer tasks, but not for verbal factual knowledge or pictorial recall tasks. The lack of an effect of type of visualization for verbal factual knowledge and pictorial recall tasks appears plausible, as the information for solving these tasks was explicitly conveyed in the multimedia instruction. However, effects of different types of visualizations are mainly expected for not explicitly taught content, where further inferences and elaborations are required from the learner, based on which a coherent mental representation of the content can be constructed (cf. Bétrancourt & Tversky, 2000). The superiority of dynamic visualizations for transfer tasks indicates that the presentation of the dynamic properties and their interrelations helped in constructing a deeper understanding of this dynamic domain. Concerning ECL, as predicted, learners presented with static visualizations found it more difficult to learn than students presented with dynamic visualizations, as indicated by the ECL_2-item and slightly by the ECL_1-item. However, there were no differences for the item supposed to measure GCL, which may be the result of a lack of validity of subjective cognitive load measures that will be discussed later (cf. de Jong, 2010).

Second, it was expected that learning with spoken as compared to written text should result in better learning outcomes, a decrease in ECL, and an increase in GCL. There was no modality effect for verbal factual knowledge tasks, but for pictorial recall tasks and transfer tasks. Because the verbal factual knowledge task could be answered based on information mainly contained in the text, one would not necessary expect a modality effect for this task, as usually text is processed dominantly and more intensively, while the processing of visualizations is subordinate (e.g., Schmidt-Weigand et al., 2010). In the same way, a possible interpretation of the superiority of spoken over written texts for pictorial recall may be that with spoken text sufficient attention could be paid to the visualizations, which contained essentially the necessary information for solving this task (see for similar results Schmidt-Weigand et al., 2010). In addition, the opportunity of processing spoken text and visualizations simultaneously rather than having to split attention between written text and pictures may have helped integrating text and pictures, thereby resulting in deeper understanding as measured in the transfer tasks. It should be noted that the modality effect was observable, even though the spoken text and the visualizations were not synchronized, in order to not favor the spoken-text over the written-text conditions, where no synchronization of text and visualizations could be realized. With respect to the items supposed to measure ECL, no modality effect could be observed. For the item supposed to measure GCL learners in the written-text condition reported to have concentrated more strongly than learners in the spoken-text conditions.

Third, contrary to the initial expectations, the modality of the text did not moderate the effectiveness of learning with either static or dynamic visualizations. Also, concerning cognitive load, no interactions between type of visualization and text modality could be observed for either the ECL or the GCL measures, respectively. It was expected that with spoken text it should be possible for learners to appropriately extract the information from the dynamic visualizations in order for their potential to unfold. It may be argued that – even though spoken text was beneficial – the visual processing demands associated with the simultaneous processing of written text and dynamic visualizations are not the main problem as compared to the demands already imposed by the dynamic visualization itself, namely the visual complexity within dynamic visualizations. Therefore, a better solution to further support learning from dynamic visualizations may involve the use of cueing techniques to guide learners' attention, which has – at least partly – been proven a useful procedure when processing dynamic visualizations (Amadieu et al., 2011; but see de Koning et al., 2011a; cf. de Koning et al., 2009 for a review). It should be noted that – with respect to the expected moderation – this experiment was designed in a conservative way, because the information in the dynamic visualizations was not transient, but repetitive. However, when using transient animations where a lack of attention is associated with a loss of information, the expected interaction may be more likely (cf. Moreno & Mayer, 1999; Schnotz, 2005). Moreover, another important point with respect to the absence of the interaction might be that the static visualizations were somewhat transient, since they were shown sequentially one after the other. Even though this format of static visualizations should rather diminish their visual complexity, it nevertheless may have intensified the drawbacks of split attention in learning with static visualizations, and hence may have overshadowed a potential moderation.

The results of the current study contradict the results of a study by Mayer et al. (2005), which calls for a closer examination. In their study, the authors found that animations and spoken text were inferior to static pictures and written text for different domains. It should be noted though that as only these two groups were compared, there was a confounding of type of visualizations and text modality. One factor, which may shed light on the contradicting results, is that in the current study, there were good reasons why dynamic visualizations should be superior to static ones, as it was about the understanding of dynamic features, whereas in the study by Mayer et al. (2005), there was no obvious reason why dynamic visualizations should be superior in these used domains (e.g., how lightning develop). However, due to the abovementioned confounding, this issue remains unclear. Moreover, the learning material used by Mayer et al. was designed in a way that it favored the written text and static pictures conditions. For instance, in Experiment 2 of their study, the auditory text could only be heard once with the animation and lasted 78 seconds (and then the animation could be replayed without auditory text until in total 6 minutes had passed), whereas the written text in the static condition was available the complete 6 minutes. But once again, due to confounding, the role of the duration of text presence cannot be elucidated. Irrespective of that fact, another important factor is that the static visualizations in the study by Mayer et al. were presented simultaneously next to each other, whereas in the current study they were shown sequentially one after the other. However, in a recent study, Boucheix and Schneider (2009) were able to show that for a mechanical content, static-simultaneous visualizations were as good for learning as dynamic visualizations, whereas static-sequential visualizations turned out to be inferior compared to both formats. As explicated in Chapter 2.2.4.2, since the current state of research concerning the effectiveness of different types of static visualizations compared to dynamic visualizations is rather inconclusive, it hence seems reasonable to investigate this issue in more detail for an ongoing study.

The covariate prerequisite knowledge had an impact on learning outcomes, in that the higher prerequisite knowledge was, the better participants achieved in the verbal factual knowledge test and the pictorial recall test, whereas prerequisite knowledge had no significant impact on transfer tasks. This means that for tasks asking for explicitly conveyed knowledge, prerequisite knowledge was helpful in achieving an understanding of the explained topics, which is in line with what the prerequisite knowledge test tested for. But for tasks asking where the learned content had to be applied to new scenarios, prerequisite knowledge was not helpful anymore, indicating that it did not affect deeper understanding, which then again was not tested in the prerequisite knowledge test. Irrespectively, it should be noted that the pattern of research results of prerequisite knowledge in learning with dynamic and static visualizations is inconsistent in itself, even when similar instructional material is used (cf. Hegarty & Kriz, 2008).

The main effect of spatial abilities on learning outcomes, specifically for tasks that require an understanding of the visualizations (like the pictorial recall tasks and transfer tasks) is often observed in multimedia research with regard to learning with dynamic and /or static visualizations (cf. Hegarty & Kriz, 2008; Höffler, 2010). However, contrary to the ability-as-compensator hypothesis, spatial abilities did not moderate learning with dynamic and static visualizations. It should be noted that even though there are good theoretical reasons, as well as meta-analytic empirical evidence (Höffler, 2010) to expect such a moderation effect, this effect is rarely found in single studies (cf. Hegarty & Kriz, 2008). As was explicated in Chapter 2.2.4.2, it might be the case that the moderating role of spatial abilities in learning with dynamic and static visualizations might particularly shine through when dynamic visualizations are compared to static-simultaneous visualizations as opposed to static-sequential visualizations. This research question will be taken up in Study 3 of this thesis, and conclusions will follow in Chapter 7 (General Discussion).

With respect to the measurement of cognitive load, the lack of an effect of modality for the items supposed to measure ECL is somewhat surprising and inconclusive, because it is neither in line with CLT nor with the replicated modality effect for learning outcomes of the current study. The observed effect for type of visualizations as well as the small, albeit significant negative correlation of the subjective ratings of both ECL-items with learning outcomes may be seen as a hint that these items at least partially measured ECL. However, the evidence for an adequate measurement of different cognitive load types by means of subjective ratings can be regarded as fragile, which renders it doubtful if different load types can be distinguished by subjective measures (e.g., de Jong, 2010). This notion also holds for the measure assumed to assess GCL, as for instance learners with written text scored higher than learners with spoken text on the GCL-item, while their performance was worse. Accordingly, in the current study, this item did not appear to have measured GCL as opposed to the study by Cierniak et al. (2009).

As the GCL-item showed a small, but positive correlation with verbal factual knowledge, but not with the other two learning outcome measures, it may rather have measured compensatory attempts, as learners stated by means of this item that they concentrated more strongly on dealing with written text than with spoken text. Independently of the condition learners were assigned to, the attempt to concentrate more strongly may have been helpful for simple verbal factual knowledge tasks, but not for a deeper understanding of the content.

Overall, even though the learning outcomes results in general were in line with CLT, this was only sparsely the case for the measures associated with ECL and GCL. If anything, the ECL_2-item (“How difficult was it for you to learn with the given material?”) seemed to be more appropriate than the ECL_1-item (“How difficult was it for you to understand the contents?”). Doubts about the adequacy and the validity of self-reports based on one-item scales that are supposed to distinguish between ECL and GCL have recently be raised, and are in line with a growing body of research using similar measures resulting in inconclusive findings (e.g., de Jong, 2010). Nevertheless, this is not to be misinterpreted in a way that the quest for finding such subjective measures should be aborted, but it simply indicates that the grail is not found yet. It also does not mean that the measures that were used in the current study should not be assessed anymore (and they will be used again in Study 3), but that it should be kept in mind that they have to be interpreted cautiously.

It should be noted though, that due to the number of changes from Study 1 to Study 2, in which most of them were supposed to favor dynamic visualizations, it is not retraceable which of these changes (or whether it was solely a conglomeration of these changes) exactly contributed to the superiority of dynamic over static visualizations with respect to transfer tasks: It might have been the improvement of the dynamic visualizations, in that dynamic features became more accentuated. But it also might be the case that learners with static visualizations may not have been able to compensate the drawbacks of static visualizations by devoting more time on them in the current study, as the learning phase was changed from self-paced to system-paced. For this latter argumentation, there is some weak indirect support from two experiments by Höffler (2007; Exp. 2 & 3): While under system-paced conditions dynamic visualizations lead to higher learning outcomes than static visualizations (Exp. 2), there were no differences between dynamic and static visualizations with regard to learning outcomes under self-paced conditions (Exp. 3), even though the same instructional material was used. Another reasonable argument may be that due to the elimination of redundancy between text and visualizations from Experiment 1 to Experiment 2, learners could less rely on the text, but had to rely more on the different types of visualizations. Moreover the impact of differences in the two types of visualizations may have been intensified, since the description of dynamic features was less extensive in Study 2. Some indirect support for this interpretation can be derived from a recent study by Kühl et al. (2010). In this study, the authors found no differences between dynamic and static visualizations when the text was rather extensive, thereby describing dynamic features in detail, whereas dynamic visualizations were superior to static visualizations, when the text was reduced, so that learners had to rely and reason more with the visualizations (for related results see also Catrambone & Seay, 2002). Moreover, it should be noted that some more changes were made from Study 1 to 2, for instance, such as no think-aloud-protocols were assessed (which may have overshadowed existent differences between dynamic and static visualizations in Study 1), or that the content of the learning material was revised and refined – and connected with this, also the knowledge tests were revised. To sum up, it is not retraceable what exactly led to the superiority of dynamic over static visualizations in Study 2, and it would need several studies to retrace it. However, the incomparableness of Study 1 and 2 was accepted to achieve the superordinate goal of improving the instructional material, and, specifically, the dynamic visualizations.

Similar to Study 2, one aim of Study 3 was to improve the instructional material, essentially the visualizations, as well as their coherence with the text. But this time, rather minor changes concerning the content of the instructional material as well as the knowledge tests were made. However, in Study 3, particularly the static-sequential visualizations were generally improved in that they were synchronized with the accompanying auditory text. In Study 2, the key frames of the static-sequential visualizations had changed every five seconds (and this in two loops), irrespective of whether the auditory text referred to the respective position that was depicted in a key frame or not. This lack of synchronization was realized to not favor the spoken text condition, as there was also no synchronization of written text and visualizations. It also should be noted that the dynamic visualizations were also not synchronized with the text.

Even though spoken text and visualizations were not synchronized, adding spoken text compared to written text led overall to better performance on the knowledge test, indicating that it had helped to overcome the problems of inter-representational split-attention, irrespective of the type of visualization. However, especially when considering dynamic visualizations, they may suffer from an overwhelming character due to a high degree of visual complexity within the visualization (Lowe, 2003, 2004). To cope with the visual complexity of dynamic visualizations, and to guide a learner's processing in a way that the benefits of dynamic visualizations might properly unfold, it has been suggested to use cueing methods (e.g., de Koning et al., 2009). Even though cueing should also be beneficial when learning with static visualizations, it was assumed that it would be even more beneficial when learners would receive dynamic visualizations (cf. Chapter 4.2). Also, as stated above, it might be worthwhile not only to investigate static-sequential visualizations, but also static-simultaneous visualizations, even though it is hardly possible to predict which format might be better suited, due to the sparse and inconsistent research in this field (cf. Chapter 2.2.4.2). However, it might be the case that static-simultaneous visualizations might benefit more from cueing than static-sequential ones, since – as explicated in Chapter 2.2.2 – static-simultaneous visualizations are supposed to be more demanding for the visual system.

In a nutshell, whereas spoken text can be regarded as a remedy to cope with inter-representational split-attention, it might not be sufficient to handle the visual complexity of dynamic visualizations themselves. To deal with the visual complexity of dynamic visualizations, and hence to optimize learning with dynamic visualizations, cueing is assumed to be helpful. The influence of cueing in learning with dynamic and static visualizations will be investigated next, thereby also taking into account two different presentation formats of static visualizations.

6 Study 3: The Impact of Cueing in Learning from Dynamic and Static Visualizations[25]

6.1 Research Question and Hypotheses of Study 3

In the previous study, it could be shown that on the one hand, spoken text lead to better performance than written text. Moreover, dynamic visualizations were more apt to convey a deeper understanding of the domain than static visualizations. Nevertheless, it still can be argued that the potential of dynamic visualizations was not completely exploited, since there still might have been a high degree of visual complexity within dynamic visualizations. Thus, to optimize learning with dynamic visualizations, and to cope with a high degree of visual complexity, it has been suggested to use cueing methods.

For cueing, a main effect for pictorial recall tasks and transfer tasks was expected, with learners in the cued conditions outperforming learners in the uncued conditions, since cueing visualizations and relating text and visualizations should mainly affect the pictorial and integrated mental model, respectively. Also, it was assumed that cueing would reduce ECL, as it should guide a learner's processing. This in turn might help learners to engage in more valuable processing activities, associated with an increase of GCL.

For type of visualization, it was expected that dynamic visualization conditions should outperform both static visualization conditions specifically for transfer tasks (but not for verbal factual knowledge or pictorial recall tasks), as differences between these types of visualizations are supposed to mainly affect tasks asking for a deeper understanding. No specified hypothesis could be derived with respect to differences between the static-sequential and the static-simultaneous conditions. Furthermore, it was expected that dynamic visualizations compared to static visualizations would result in a reduction of ECL. However, since there was no impact of type of visualization on GCL in Study 2 of this thesis, no hypothesis was formulated concerning this topic.

Furthermore, an interaction between type of visualization and cueing was hypothesized: Dynamic visualizations might benefit comparatively more from cueing than any presentation format of static visualizations, since static visualizations are supposed to possess a lower degree of visual complexity. Concerning static visualizations, the benefits of cueing should be more

[25] This chapter is based on: Kühl, T., Scheiter, K., & Gerjets, P. (2011). *The impact of cueing in learning with dynamic and static visualizations.* Manuscript in preparation.

pronounced for static-simultaneous visualizations than for static-sequential visualizations, since the attentional demands on static-simultaneous visualizations might be higher than those of static-sequential visualizations, so that static-simultaneous visualizations might benefit more from cueing than static-sequential ones. This pattern of results should be mirrored by the respective pattern of ECL and GCL.

Moreover, it was once again examined, if spatial abilities would moderate learning with dynamic and static visualizations as predicted by the ability-as-compensator hypothesis: Especially learners with weaker spatial abilities should profit from dynamic visualizations, whereas the benefit of dynamic visualizations should be less pronounced for learners with stronger spatial abilities. Even though this moderating role of spatial abilities was not evident in Study 2, this research question was kept, since in contrast to Study 2, in Study 3 additionally static-simultaneous visualizations were used. However, one may argue that the moderating role of spatial abilities might become evident, when comparing dynamic visualizations with static-simultaneous visualizations, as especially for the latter ones mentally animating the movements becomes more important. This potential moderating effect should also hold for the items supposed to measure cognitive load.

6.2 Method

6.2.1 Participants and Design

One hundred and fifty students (122 female and 28 male participants; M = 22.47 years, SD = 3.07) with various educational backgrounds from the University of Tuebingen, Germany, participated either for course credit or payment in the study. Students had to be native speakers of German; no students of physics were allowed to take part. Students were randomly assigned to one of six conditions, which resulted from a 2x3-design with cueing (with/without) and type of visualization (dynamic, static-sequential, static-simultaneous) as independent variables. Twenty-five participants served in each condition.

6.2.2 Instructional Materials

The computerized learning material dealt with the physical principles underlying fish locomotion. This topic addresses the understanding of physical concepts in relation to movement characteristics such as trajectory, velocity, and acceleration. The material consisted of eight sections, which built upon each other. Particularly, these sections contained the themes 1) Swim styles, 2) Pushing off the water, 3) Body section and propelling element, 4) Undulatory motion, 5) Actio and Reactio, 6) Magnitude of the reaction force, 7) Decomposing the reaction force, as well as 8) Interaction of forces of various propelling elements. Note that in comparison to Study 2 of this thesis, the section "2) Pushing off the water" was added to the instructional material of Study 3. Additionally, some minor adjustments were realized. These changes aimed at making the instructional material more comprehensive.

Each section consisted of visualizations and corresponding explanatory texts. The same spoken text (695 words) was used in all conditions. The learning material was presented system-paced. Each section lasted between 45 to 77 seconds (in total 481 seconds), corresponding to the length of the spoken text for each section.

The visualizations were subject to experimental manipulation and differed with regard to type of visualization (dynamic, static-sequential, static-simultaneous) and the presence of cueing (with/without). Irrespective of these manipulations, all visualizations were placed in the middle of the screen.

In conditions with dynamic visualizations, there was always one animation showing an undulatory (i.e., wave-like) movement of a fish in a recursive fashion (see Figure 6.1). The animation depicted the same fish across the eight instructional sections, but focused on different aspects of its movement by portraying the interplay of the trajectory and velocity of different body parts, the corresponding displacement of water, the sizes of the associated resulting forces and their direction, as well as the related swimming speed. These forces were represented as arrows and varied in length and spatial orientation depending on the force's strength and direction. For instance, in the section explaining the magnitude of the reaction force, the fish changed its frequency of the movement of the body parts to depict the relation of these changes and the associated changes in the sizes of the resulting forces (i.e., changing size of arrows) and the related swimming speed (i.e., changing speed of moving background).

In conditions with static-sequential visualizations, nine key frames were shown within each section that had been extracted from the corresponding animation. The key frames were displayed sequentially one after the other. The nine static key frames represented two loops of an undulatory movement, so that each learner had the chance to see a frame again in case he/she had missed the information the first time (see Figure 6.1). Each key frame of a section remained visible on the screen between five up to 14.5 seconds, depending on the time the spoken text referred to the particular position of the fish. As can be seen in Figure 6.1, key frame number one, number five and number nine showed the same position, as they represented the starting point and the end point of the undulatory fish movement. Moreover, always two key frames showed identical positions of the fish (i.e., key frames two and six, key frames three and seven, as well as key frames four and eight).

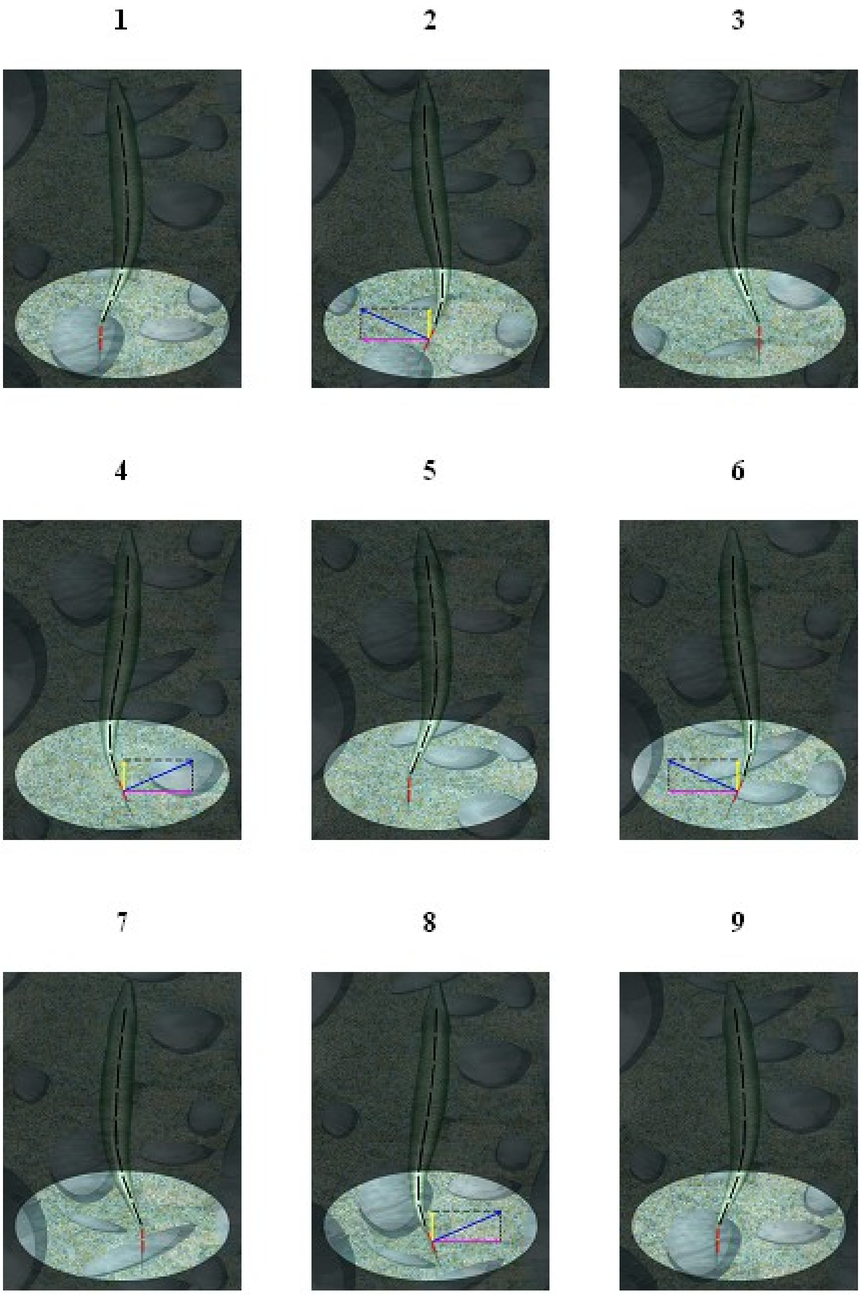


Figure 6.1. Sequence of nine static key frames of the cued static-sequential visualization condition, depicting the spotlight and different elements in different colors. Note that the arrows were no cueing device, but belonged to the content and were used to symbolize the forces resulting from the fish's movement.

In conditions with static-simultaneous visualization, the first five key frames of the static-sequential visualizations were shown within each section[26]. They were presented next to each other, ordered from left to right. Each key frame of the simultaneous condition had only a fifth of the area of a key frame of the static-sequential visualization condition, so that they would fit on the monitor screen. Irrespective of the smaller size, everything relevant could be seen in the respective key frames.

The cueing procedure for visualizations consisted of different cueing methods: (a) showing elements only after they were mentioned in the text, (b) color-coding, (c) a spotlight, as well as (d) zoom-ins (and zoom-outs). These methods were implemented for all types of visualizations, with the aim to counteract the assumed visual complexity of dynamic visualizations. Additionally, these cues were designed to support the processes of selecting, organizing and integrating information. In the following, the different functions of these cues will be described in more detail.

First, to not to interfere with the depicted movement, a spotlight and colored elements were implemented to emphasize important aspects (see Figure 6.1), thereby producing a visuo-spatial contrast (e.g., instead of using arrows as a cueing device) and supporting the process of selecting.

Second, the spotlight was also applied to overshadow less relevant and potentially distracting movements. In a spotlight, the most relevant parts are in its focus, whereas the distracting movement of the background is reduced because it is deemphasized. Note, that the moving background was most of the time not irrelevant, because it depicted dynamic information, namely the actual swimming speed of the fish. Therefore, it did not seem advisable to leave out this information. However, as the moving background was covering the complete animation, it might have been more distracting than necessary. By applying a spotlight, a part of the background in the focus was still entirely visible, whereas the rest of the moving background was dimly visible only. As a consequence, the swimming speed was still perceivable, whereas the background was supposed to be less distracting. For depicting the changes of the velocity of the caudal fin from reversal point to baseline to reversal point – when describing the undulatory movement and the magnitude of the reaction force – a moving background was not considered necessary, but if anything disturbing. In this case, the background was completely eliminated and it was zoomed in to the caudal fin (and later zoomed out again). Note that for each section in which a spotlight was realized, it appeared only after the visualization had been present for 2

[26] Note that it was not necessary to show key frames number six to nine in the static-simultaneous visualizations condition as compared to the static-sequential visualizations conditions, since for the static-sequential visualization conditions these key frames were shown only to decrease the chances of missing information.

seconds, so that participants could realize that the spotlight served as a manipulation within the visualization (e.g., de Koning et al., 2010a). Overall, the spotlight was supposed to reduce the visual complexity, and to aid learners to select the most relevant information.

Third, dynamic relations were overemphasized, particularly the changes of the velocity of the caudal fin from reversal point to baseline to reversal point and its impact on the reaction force. This was done by overemphasizing the acceleration and deceleration of the movement pattern of the caudal fin (which was accompanied by removing the background and zooming in; see above). Moreover, when it seemed reasonable, the dynamic visualizations stopped at crucial states to make this information better extractable for learners. For instance this was done to emphasize an important position of the fish's tail and its impact on the direction of the corresponding forces to highlight important aspects of this state. Note that these two cueing methods, namely overemphasizing and stopping, could uniquely be implemented for dynamic visualizations, and aimed at helping learners to more easily select the depicted information.

Fourth, to correspond to the functional aspects of a system, elements were depicted according to their functional aspects, for instance, all depicted propelling forces were colored yellow, all reaction forces were colored blue, and all lateral forces were colored violet (see Figure 6.2). Moreover, elements were presented according to the cause-and-effect chain (e.g., first the displaced water, thereafter the resulting reaction force). At this, it was supposed that these methods would support learners to organize the information into a coherent pictorial mental model.

Finally, cues were designed to emphasize the relationship of elements of the text and the visualization, with the intention to make it easier for a learner to integrate the text and visualizations into a coherent integrated mental model (e.g., by not depicting elements in the visualizations until they were mentioned in the text, or by coloring elements in the visualization when mentioned in the text). By doing so, also the demands to split the attention within the visualizations were reduced, since the visualizations were build up in a stepwise manner by introducing relevant elements one after another (see above).

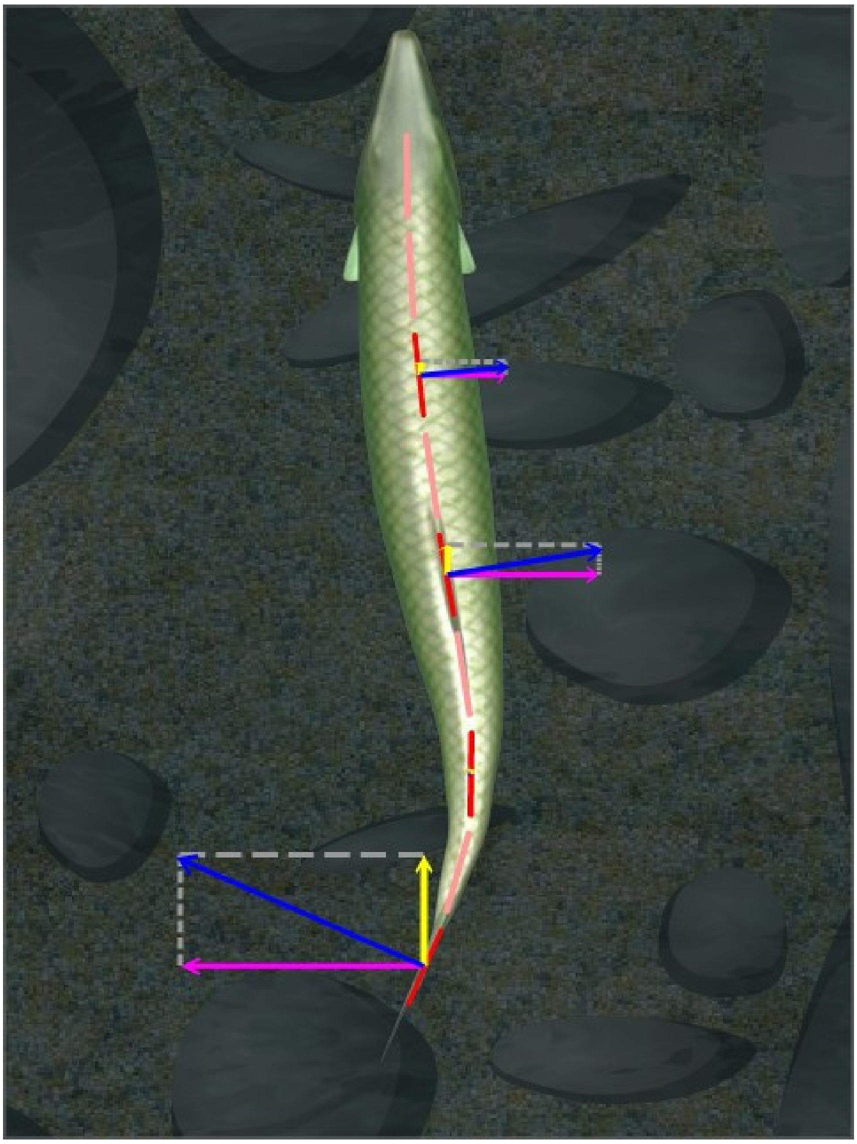


Figure 6.2. Snapshot of a cued visualization condition depicting four parallelogram of forces, at which the same type of forces receive the same color. Note that for this case no spotlight was implemented, since several spots were equally relevant, but that the background was dimmed.

6.2.3 Measures

A questionnaire with respect to the attitudes towards biology and physics, as well as a prerequisite knowledge test served as control variables, and a spatial ability test served as a moderator variable (and also as a control variable). As dependent variables served on the one hand two items asking about the experienced cognitive load during the learning phase, and furthermore several knowledge tests to measure learning outcomes.

Attitudes towards biology and physics. Concerning the attitude towards biology and physics, the same questionnaire as in Study 2 was used. It consisted of ten items, which had to be rated on a 4-point Likert scale. Note that negatively formulated items were recoded. Five items dealt with the attitude towards biology and were subsumed to one score, while the other five items dealt with the attitude towards physics and were subsumed to another score. Their internal consistencies in this study can be classified as very good, with $\alpha = .87$ of the biology scale, and $\alpha = .89$ of the physics scale.

Prerequisite knowledge. The same prerequisite knowledge test as in Study 2 was used. The test consisted of eight questions asking for the second and third Newton axioms, the physical definition of forces, the characteristics of a harmonic oscillation, knowledge about velocity and acceleration, knowledge about the undulatory swimming style, and knowledge about the interplay of forces by analyzing a parallelogram of forces (see below for a sample item). A person's knowledge about these basic definitions and principles was considered a beneficial prerequisite for more easily achieving an understanding of the topics explained in the current study. The eight multiple-choice questions consisted of four to six alternatives to choose from and for each question there were one to three correct answers. For each correct answer, learners were assigned one point and for each wrong answer one point was subtracted. Within a question, however, learners could at worst receive zero points. The maximum score was 12 points.

Example of a question from the prerequisite knowledge test

According to Newton's second law of motion, a force *F* is calculated from

a) the product of mass and time

b) the product of mass and acceleration

c) the product of time and impulse

d) the product of impulse and acceleration

Spatial ability. To control for individual differences in spatial abilities and to examining their potential moderating role, the mental rotation test (MRT) was administered (Vandenberg & Kuse, 1978). The MRT was used, as especially learning with the static visualizations used in this study required the ability to mentally rotate and manipulate visuo-spatial objects (e.g., to imagine the movement of the caudal fin). The MRT consists of 20 items, whereby each item comprises a complex three-dimensional block figure and four alternative figures as multiple-choice answer options. For each item, the participant has to choose, which two of the four alternative figures are identical to the target when (mentally) rotated. There is a time limit of six minutes for working on the MRT. For each correctly identified figure one point was given and for each wrong identified figure one point was subtracted, resulting in a maximum of 40 points and a minimum of -40 points.

Cognitive load measures. Since the results of Study 2 indicated that the ECL_2-item ("How difficult was it for you to learn with the given material?") seemed to be more appropriate than the ECL_1-item ("How difficult was it for you to understand the contents?"), the ECL_2-item from Study 2 was used in Study 3. Moreover, the item supposed to measure GCL from Study 2 ("How much did you concentrate during learning?") was again assessed in Study 3. These two items had to be rated on a nine-point Likert scale.

Knowledge tests. Learning outcomes were measured by means of verbal factual knowledge tasks (eleven multiple-choice questions), five pictorial recall and twelve transfer tasks (see below for sample items of each test). A maximum of 21 points could be achieved for the verbal factual knowledge test, a maximum of 9 points could be achieved for the pictorial recall test, and a maximum of 32.5 points could be achieved for the transfer test. The verbal factual knowledge tasks were posed in a verbal format and all correct answers were explicitly conveyed by the multimedia instruction. The pictorial recall tasks were posed in a pictorial format and asked for facts that were depicted by the visualizations. The transfer tasks were posed in written as well as in pictorial form. For solving the transfer tasks, learners had to apply their acquired knowledge to new situations.

The open questions of the pictorial tasks (4 of 5) as well as the transfer tasks (6 of 12) of 30 participants were scored by two independent raters. Cases of disagreement (3.26%) for the open questions were resolved by consensus. As interrater-reliability was high (Cohen's kappa: .95), the remaining data were scored by one rater. Performance was transformed into percentage correct for ease of interpretation.

Example of a question from the verbal factual knowledge test

Which of the following is/are true?

a) The reaction force is perpendicular to the propelling element.

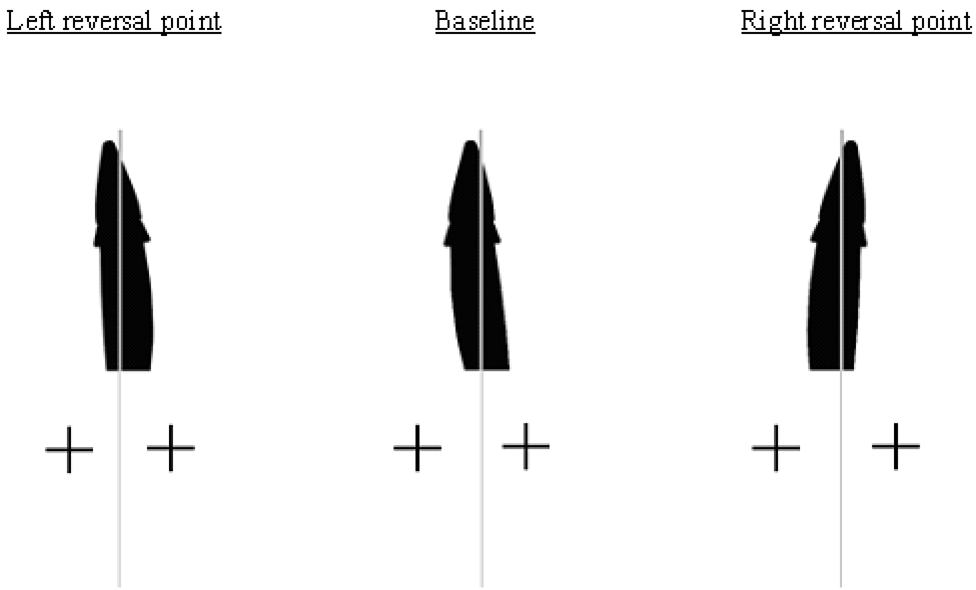
b) Lateral force and reaction force are perpendicular to each other.

c) Propelling force and lateral force are perpendicular to each other.

d) Reaction force and propelling force are perpendicular to each other.

Example of a question from the pictorial recall test

Three positions of an undulatory swimming fish are given below. The grey line symbolizes the baseline and the crosses symbolize the reversal points of the tail. However, the lower part of the fish is covered. Please draw the lower parts of the fish for the three given positions.



Example of a question from the transfer test

Some undulating species of fish move their head back and forth in order to swim forwards. Why is this? Write down any feasible reasons you can think of!

6.2.4 Procedure

Each participant was tested individually in a session lasting up to 120 minutes. First the different control variables were assessed in the following order: attitudes towards biology and physics, the prerequisite knowledge test, and the MRT. Thereafter, the learning phase began and when this phase was finished, students had to rate their cognitive load experienced during learning. Then students had to work on the different knowledge tests, namely the verbal factual knowledge test, followed by the pictorial recall test and finally the transfer test. With the exception of the MRT, which was paper-pencil based, all other materials were presented via computer.

6.3 Results

Means and standard deviations are reported in Table 6.1. Partial eta-squared ($\eta^2 p$) is reported as measures of effect size.

Table 6.1

Means (and SD) as a Function of Type of Visualization and Cueing

Type of visualization	Dynamic		Static-sequential		Static-simultaneous	
Cueing	Yes (*n* = 25)	No (*n* = 25)	Yes (*n* = 25)	No (*n* = 25)	Yes (*n* = 25)	No (*n* = 25)
Control variables						
Attitudes towards biology (5 – 20)	16.88 (2.91)	17.80 (2.63)	17.12 (3.27)	17.16 (3.46)	18.56 (2.76)	16.92 (3.73)
Attitudes towards physics (5 – 20)	10.72 (2.81)	12.68 (2.93)	12.80 (3.65)	13.20 (3.77)	12.36 (3.66)	13.24 (4.05)
Prior knowledge (% correct)	49.67 (15.12)	59.33 (19.29)	54.67 (19.85)	52.33 (18.24)	50.67 (19.97)	57.67 (21.51)
Spatial abilities (-40 – 40)	17.28 (10.26)	18.36 (8.22)	19.72 (10.13)	20.52 (9.77)	23.68 (9.43)	21.00 (9.70)
Learning outcomes (% correct)*						
Verbal factual knowledge	70.48 (3.54)	66.71 (3.43)	66.52 (3.40)	59.72 (3.41)	60.94 (3.53)	62.96 (3.42)
Pictorial recall	53.03 (4.39)	45.42 (4.26)	45.96 (4.22)	39.92 (4.23)	54.11 (4.37)	41.49 (4.24)
Transfer	51.63 (2.81)	48.59 (2.73)	46.28 (2.71)	45.39 (2.71)	45.28 (2.80)	43.14 (2.72)
Cognitive load (1-9)*						
ECL	3.45 (.34)	3.93 (.33)	4.62 (.32)	4.49 (.33)	5.10 (.34)	5.33 (.33)
GCL	6.69 (.36)	6.00 (.35)	6.58 (.34)	5.93 (.35)	5.88 (.36)	5.72 (.35)

* Learning outcomes are adjusted by taking into account attitude towards physics and z-standardized scores for spatial abilities as covariates; values in parentheses refer to standard errors for these dependent measures.

6.3.1 Comparability of Experimental Conditions with Respect to Attitude Towards Biology and Physics, Prerequisite Knowledge, and Spatial Abilities

A two-factorial ANOVAs with cueing (with/without) and type of visualization (dynamic/static-sequential/static-simultaneous) was conducted to analyze if the learners in the six experimental conditions possessed similar prerequisite knowledge, spatial abilities, and attitudes towards biology as well as physics. Concerning prerequisite knowledge, there were no differences for cueing ($F(1, 144) = 2.35$, $MSE = 364.82$, $p = .13$, $\eta^2 p = .02$), type of visualization ($F < 1$, *ns*), and no interaction ($F(2, 144) = 1.36$, $MSE = 364.82$, $p = .26$, $\eta^2 p = .02$). With regard to spatial abilities, there were no differences for cueing ($F < 1$, *ns*), and no interaction ($F < 1$, *ns*), but marginal significant differences occurred for type of visualization ($F(2, 144) = 2.77$, $MSE = 29.23$, $p = .07$, $\eta^2 p = .04$). For attitude towards biology, there also were no differences for cueing ($F < 1$, *ns*), or type of visualization ($F < 1$, *ns*), and no interaction ($F(2, 144) = 2.13$, $MSE = 9.92$, $p = .12$, $\eta^2 p = .03$). For attitude towards physics, there were no differences for type of visualization ($F(2, 144) = 1.99$, $MSE = 12.31$, $p = .14$, $\eta^2 p = .03$), and no interaction ($F < 1$). However, there was a marginal significant difference for cueing ($F(1, 144) = 3.55$, $MSE = 12.31$, $p = .06$, $\eta^2 p = .03$), with students in the uncued conditions ($M = 13.04$, $SD = 3.58$) possessing more positive attitudes towards physics than students in the cued conditions ($M = 11.96$, $SD = 3.47$). Because the experimental conditions could not be regarded as equal with respect to learners' attitudes towards physics, this variable was used as covariate in all analyses reported in this Study. Since spatial ability is inserted in the analyses as a continuous factor to test its moderating role, differences among the visualization conditions are controlled for.

6.3.2 Effects of Cueing, Visualization Format, and Spatial Abilities on Learning Outcomes

Two-factorial ANCOVAs with cueing and type of visualization as independent variables and verbal factual knowledge, pictorial recall and transfer, respectively, as dependent variables were conducted, with learners' attitudes towards physics as covariate. Furthermore, to test if spatial abilities moderated learning with the three types of visualizations, the scores of the MRT were z-standardized and their interaction with type of visualization was inserted into the respective ANCOVAs.

With respect to type of visualization, the two-factorial ANCOVAs revealed no differences for either verbal factual knowledge ($F(2, 140) = 2.04$, $MSE = 288.64$, $p = .13$, $\eta^2 p = .03$), or pictorial recall (F(2, 140) = 1.18, MSE = 444.34, p = .31, $\eta^2 p$ =.02). For transfer, there was a marginally significant effect for type of visualization (F(2, 140) = 2.39, MSE = 182.68, p = .097, $\eta^2 p$ =.03). Because it was hypothesized that dynamic visualizations would outperform both types of static visualizations for transfer tasks, planned contrasts between the dynamic visualization conditions and the two types of static visualization conditions were applied, using attitude towards physics as covariate and spatial abilities as continuous factor. Results revealed a significant effect (F(1, 140) = 4.42, MSE = 182.68, p = .04, $\eta^2 p$ =.03), indicating that, in line with the hypothesis, learners in the dynamic visualization conditions outperformed learners in both static visualization conditions for transfer tasks. There were no significant differences between the static-sequential and static-simultaneous conditions for transfer tasks (F < 1, *ns*).

The two-factorial ANCOVAs revealed no effect of cueing on verbal factual knowledge (F(1, 140) = 1.02, MSE = 288.64, p = .32, $\eta^2 p$ =.01) and, in contrast to the hypotheses, also no effect of cueing on transfer (F < 1, *ns*). However, cueing had an effect on pictorial recall (F(1, 140) = 6.25, MSE = 444.34, p = .01, $\eta^2 p$ =.04), with learners in the cued conditions (M = 51.03, SE = 2.50) outperforming learners in the uncued conditions (M = 42.28, SE = 2.46).

With respect to the assumed interaction between cueing and type of visualization, the two factorial ANCOVAs revealed no interaction for any of the three learning outcome measures (all *Fs* < 1, *ns*).

Spatial abilities had an impact on all learning outcome measures: on verbal factual knowledge (F(1, 140) = 8.70, MSE = 288.64, p < .01, $\eta^2 p$ =.06), on pictorial recall (F(1, 140) = 8.00, MSE = 444.34, p < .01, $\eta^2 p$ =.05), and on transfer (F(1, 140) = 21.00, MSE = 182.68, p < .001, $\eta^2 p$ =.13). The higher spatial abilities were, the better was performance on verbal factual knowledge tasks (r = .23, p < .01), on pictorial recall (r = .26, p = .001), as well as on transfer (r = .37, p < .001). However, other than expected, there was no interaction of spatial abilities and type of visualizations for pictorial recall (F(2, 140) = 1.04, p = .36, $\eta^2 p$ =.02), verbal factual knowledge, or transfer (both *Fs* < 1, *ns*), indicating that spatial abilities did not moderate learning with the different types of visualizations.

6.3.3 Effects of Cueing, Visualization Format, and Spatial Abilities on Cognitive Load

For each cognitive load item a separate two-factorial ANOCVA was conducted with cueing and type of visualization as independent variables, attitude towards physics as covariate, and the ECL-item or GCL-item, respectively, as dependent variable. Like for learning outcomes, it was furthermore tested if spatial abilities moderated experienced cognitive load when learning with the three types of visualizations by introducing it as a continuous factor in the analyses.

Concerning type of visualization, the two-factorial ANCOVA revealed no main effect for the GCL-item ($F(2, 140) = 1.29$, $MSE = 3.06$, $p = .28$, $\eta^2 p = .02$), but for the ECL-item as had been expected ($F(2, 140) = 10.25$, $MSE = 2.66$, $p < .001$, $\eta^2 p = .13$). Concerning the ECL-item, planned contrasts between the dynamic and the two types of static visualizations conditions revealed a significant effect ($F(1, 140) = 16.53$, $MSE = 2.66$, $p < .001$, $\eta^2 p = .11$), with learners in dynamic visualization conditions perceiving the content as less difficult than learners in the two types of static visualizations condition. Moreover, there was a significant effect between the static-sequential and the static-simultaneous visualization conditions ($F(1, 140) = 4.13$, $MSE = 2.66$, $p = .04$, $\eta^2 p = .03$), with participants in the static-sequential visualization conditions perceiving it as less difficult to learn with the content than learners with static-simultaneous visualizations.

Cueing had no effect on the ECL-item ($F < 1$, *ns*), but a marginal main effect on the GCL-item ($F(1, 140) = 2.87$, $p = .09$, $\eta^2 p = .02$), with learners in the cued conditions ($M = 6.37$, $SE = .21$) reporting slightly more concentration on the content than learners in the uncued conditions ($M = 5.88$, $SD = .20$).

There was no interaction between type of visualization and cueing, neither for difficulty, nor for concentration (both $Fs < 1$, *ns*).

Spatial abilities had no effect on the ECL-item ($F(1, 140) = 2.22$, $MSE = 2.66$, $p = .14$, $\eta^2 p = .02$) or on the GCL-item ($F < 1$, *ns*). The two-factorial ANCOVAs did not reveal an interaction of spatial abilities and type of visualization for the ECL-item or GCL-item (both $Fs < 1$, *ns*).

The items supposed to measure ECL and GCL did correlate significantly with each other ($r = -.30$, $p < .001$) in that higher ECL ratings were associated with lower GCL ratings and vice versa. To analyze if the subjective cognitive load ratings were related to performance, correlations were conducted between the ECL and GCL measures and verbal factual knowledge, pictorial recall, and transfer (see Table 6.2). The three learning outcome measures were always negatively associated with ECL, and always positively associated with GCL. These correlations are in line with what would be expected from CLT and might be considered as a positive validation check for these items.

Table 6.2

Correlations Among Cognitive Load Measures and Knowledge Tests

n = 150	Verbal factual knowledge	Pictorial recall	Transfer knowledge
ECL	*r* = -.27**	*r* = -.35***	*r* = -.31***
GCL	*r* = .25**	*r* = .31***	*r* = .18*

Note: *** p < .001; ** p < .01; * p < .05

6.4 Summary and Discussion

In the current study, it was investigated if learning with multimedia instruction in general, and learning with dynamic visualizations in particular, could be optimized by means of cueing.

First, it was expected that learning with dynamic as opposed to static visualizations would result in better learning outcomes, specifically for transfer tasks, and a decrease in ECL. The findings of this study supported this assumption to a large extent. On the one hand, learners in the dynamic visualizations conditions outperformed learners in the different static visualization conditions for transfer tasks, whereas no significant differences occurred for verbal factual knowledge and pictorial tasks, thereby confirming the results of Study 2. Concerning the two formats of static visualizations, there were no differences for any of the three learning outcome measures, indicating that the format of the static visualizations played a subordinate role for this kind of learning material. The superiority of dynamic visualizations for transfer tasks indicates that the presentation of dynamic features like changes in velocity and their interrelations helped in constructing a deeper understanding of this domain. For future studies, it would be interesting to check the generalizability of this finding to other domains with similar dynamic features, for instance Kepler's second law, in which the understanding of the changes in velocity of the planetary motion around the sun is crucial (cf. Kühl et al., 2010). The item supposed to measure ECL correlated negatively with all learning outcomes measures, which may be regarded as a positive validation check. As predicted for ECL, learners presented with static visualizations found it more difficult to learn the content than students presented with dynamic visualizations. Moreover, participants in the static-simultaneous conditions found it more difficult to learn than participants in the static-sequential visualizations conditions. The item assessing learner's concentration was supposed to be measure for GCL, and the positive relationship of this item with the learning outcome measures may support this assumption. For this item, similarly as in Study 2 of this thesis, no differences with regard to type of visualization could be observed. This finding may be attributed to a minor influence of the type of visualization on GCL, provided this item indeed measured GCL. It should be noted that even though the correlations of the ECL and GCL item with the learning outcome measures are in line with what would be expected from CLT, nevertheless the interpretation of these items should be treated cautiously, since doubts are raised from time to time whether subjective cognitive load ratings can be regarded as a valid measurement of cognitive load (cf., de Jong, 2010).

Second, it was expected that learners in the cued conditions would outperform learners in the uncued conditions, particularly for pictorial recall and transfer tasks, and that cueing would result in a decrease of ECL as well as in an increase of GCL. As expected, there was an effect of cueing for pictorial recall tasks, and not for verbal factual knowledge tasks. However, other than expected, cueing had no influence on transfer tasks. Hence, on the one hand, cueing helped learners to better recall the information depicted in the visualizations, but this did not lead to a deeper understanding of the content. One possible explanation may be that cueing mainly helps in mentally organizing and structuring visual elements, but contributes less to a more elaborated mental model from which inferences concerning interrelations of dynamic features can be drawn. This finding partly corresponds to a recent review by de Koning et al. (2009), who could only find a positive effect of cueing in animations for approximately half of the reviewed studies and concluded that cueing does not necessarily lead to a deeper understanding of the content. Nevertheless, this finding is unexpected and somewhat discouraging, since the implemented cues in the current study were designed in a careful way, thereby incorporating successful cueing methods from other studies that aimed at supporting the processes of selecting, organizing and integrating information, and that, furthermore, aimed at reducing the visual complexity within (dynamic) visualizations. With respect to the item supposed to measure ECL, no effect of cueing could be found. On the one hand, this contradicts the abovementioned assumption, but on the other hand, it also mirrors the effect of cueing on transfer tasks, so this pattern of results is at least consistent in itself. For the item supposed to measure GCL, learners in the cued conditions gave higher ratings than learners in the uncued conditions. While this does not reflect the performance on transfer tasks, it is well in line with performance on pictorial recall tasks. One may interpret this finding in that cueing may help to concentrate on the content, as cueing is supposed to guide a learner's processing.

It should be noted though that cueing consisted of different cueing methods, such as synchronizing text and visualizations, a spotlight, zooming et cetera. All these manipulations were assumed to have a positive effect, and, hence were implemented to improve the instructional material. Consequently, it cannot be traced back which impact each manipulation had on the learner's processing of the materials. Also, it cannot be completely ruled out that different cueing manipulations interfered with each other - even if that appears to be unlikely, because prior studies indicated that the use of multiple cueing techniques is associated with best performance (e.g., Jamet et al., 2008). Nevertheless, the unique contribution of each cueing technique and their potential interactions need to be examined in future studies.

Third, other than expected, cueing did not moderate the effectiveness of learning with either dynamic or any format of static visualizations. Also, concerning cognitive load, no interactions between type of visualization and cueing could be observed for either ECL or GCL. On the one hand, it was assumed that cueing should help to reduce the visual complexity and to guide a learner's processing, which should be especially beneficial in dynamic visualizations, because dynamic visualizations are supposed to possess a comparatively high degree of visual complexity (cf. Lowe, 2003). Moreover, some cueing methods could uniquely be implemented for dynamic visualizations (e.g., overemphasizing changes in velocity), thereby, if anything, fostering the supposed interaction. However, the assumption of a moderating role of cueing in learning with dynamic and static visualizations was not confirmed. While cueing had no impact on transfer tasks for any type of visualization, it had an equally positive influence on pictorial tasks for all types of visualizations. Based on these results, several explanations might account for this finding. On the one hand, it may be argued that because several cueing techniques were applied, they may have had different influences on the respective types of visualizations. For instance, it may be possible that the manipulation of zooming in the visualizations was most helpful for the static-simultaneous visualizations, because the key frames shown in this condition were presented in the smallest size, whereas adding elements to the visualization only once they have been mentioned in the narration might have been most advantageous for learning with dynamic visualizations. Even though this explanation is notional, from a more principle-based point of view it might be worthwhile to investigate each manipulation separately. Another explanation for the absence of a moderating influence of cueing in learning with dynamic and static visualizations might be that the visual complexity of the dynamic visualizations for this specific learning material might not have been the major problem. This may be the case, because, first, the content was segmented into multiple sections (cf. segmenting principle, Mayer, 2009), where the sections built up on each other and, correspondingly, the number of elements in the visualizations increased from section to section. Secondly, dynamic visualizations were shown repeatedly, so that learners in the uncued conditions had the chance to see relevant changes several times. These two factors might already have decreased the visual complexity of dynamic visualizations to a certain degree, and thus may have overshadowed additional potential cueing effects for dynamic visualizations, for which the strongest effect for cueing was expected. Hence, one may speculate that a moderating role of cueing in learning with dynamic and static visualizations might be observed, when the animations are for instance transient, so that a lack of or misdirected attention will be associated with a loss of information.

Contrary to the advocated spatial ability-as-compensator hypothesis, no moderating effect of spatial ability in learning with dynamic and static visualizations could be observed, irrespective of the presentation format of static visualizations (i.e., static-sequential and static-simultaneous visualizations, respectively). It should be noted though that the moderating role of spatial ability is rarely found in single studies (cf. Hegarty & Kriz, 2008), so that the specific circumstances for the moderating role of spatial ability should be investigated in more detail in ongoing studies. Such circumstances might be the used test to measure spatial abilities, or to which factor of spatial ability the used test may contribute, respectively (cf. Höffler, 2010), and how the applied tests matches with the task to mentally animate the changes, when receiving static visualizations. Nevertheless, spatial ability had a strong effect on learning outcomes in that higher spatial abilities were associated with higher learning outcomes, a finding that is predominantly observed when learning with visualizations (cf. Hegarty & Kriz, 2008; Höffler, 2010).

To sum up, the results of this study confirmed the hypothesis that dynamic visualizations in contrast to the two types of static visualizations were more apt for learners to get a deeper understanding of a dynamic content like the one at hand. Cueing on the other hand, had no influence on tasks asking for a deeper understanding of the content, but it generally helped learners to better recall the information depicted in the visualizations. These results were – at least partly – mirrored by the items that were supposed to measure cognitive load. No moderating influences could be observed for learning with dynamic compared to static visualizations, that is neither the design factor cueing nor the learner characteristic spatial ability moderated learning with dynamic and static visualizations. It remains unclear from this study, if these missing moderating effects can, for instance, be attributed to the design of the instructional material (e.g., the repetitive movement, or the segmented instructional material). Hence, future studies using different materials may investigate these questions in more detail. All in all, from an instructional point of view, these results indicate that it is worthwhile to produce dynamic visualizations for learning domains with properties like the one at hand, where the understanding of dynamic features like changes in velocity is crucial.

7 General Discussion

7.1 Aims and Research Questions

In the current thesis, it was examined how to foster the comprehension of Natural Science phenomena by means of deploying multimedia instruction. Thereby, the domain physical principles underlying fish locomotion was chosen, for at least two reasons: First, the domain can be regarded as a concrete solution to deal with the problems of decontextualization that learners are often confronted with in learning Natural Sciences, and particularly in learning physics (cf. Taasoobshirazi & Carr, 2008; Whitelegg & Edwards, 2002). More precisely, the physical principles underlying fish locomotion are a concrete instantiation of the abstract Newton's laws of motion with reference to the real world (Waltner, Rachel, et al., 2006; Waltner, Wiesner, et al., 2007). Second, the chosen domain mirrors the requirements learners are often confronted with in the Natural Sciences, in particular, learners need to understand how a change in one variable affects another variable, in this case, for instance, the interplay of the changing velocity of a fish's caudal fin along its trajectory, and its impact on the sizes of the associated resulting forces, or its impact on the related swimming speed.

Based on a literature review it was proposed that adding visualizations to text would be helpful in fostering students' understanding of Natural Science phenomena, such as the physical principles underlying fish locomotion. In particular, dynamic visualizations as opposed to static visualizations were assumed to possess enormous potential to convey dynamic interrelations, like changes in velocity and their impact on further variables. Therefore, in Study 1 it was investigated whether static and dynamic visualizations would foster learning compared to a pure text-based instruction, as well as whether dynamic visualizations would be even better suited to achieve this aim. Based on the literature concerning the cognitive functions dynamic and static visualizations may play for learning as opposed to only text, students' cognitive processes were investigated by means of verbal protocols

Even though in particular dynamic visualizations may be well suited for conveying the physical principles underlying fish locomotion, their processing might be hampered due to their high degree of visual complexity. Therefore, the goal of Study 2 and 3 of this thesis was to investigate ways of optimizing learning with multimedia in general, and learning with dynamic visualizations in particular, so that the potentials of dynamic visualizations might best unfold. It was assumed that due to the higher degree of visual complexity in dynamic visualizations compared to static visualizations, the effectiveness of dynamic visualizations might be reduced under inter-representational split-attention conditions, as is for instance the case when presenting written instructional explanations along with the visualizations. Hence, in Study 2, it was investigated if the superiority of dynamic over static visualizations would be more pronounced when using spoken instead of written text. On the other hand, it was assumed that due to the higher degree of visual complexity of the dynamic visualizations compared to static visualizations, the selection and organization of relevant information in the dynamic visualizations, as well as its integration with text, would be particularly challenging. Hence, in Study 3, it was examined if the superiority of dynamic over static visualizations could be even enhanced by means of cueing. Moreover, in all three studies it was examined if the benefits of dynamic over static visualizations would especially be observable for learners with weaker spatial abilities, as would be predicted by the ability-as-compensator hypothesis (cf. Mayer & Sims, 1994).

In the following, first the findings of the three studies will be summarized, and afterwards each research question will be taken up again. Finally, limitations of the studies of the current thesis will be discussed and conclusions will be given.

7.2 Summary of Main Findings

In Study 1 of this thesis, one major goal was to test if adding visualizations to text would generally help learners to better understand the chosen domain in this thesis. Furthermore, the assumption was examined whether dynamic visualizations would be better suited than static visualizations to convey a deeper understanding for this domain. At this, and in line with the ability-as-compensator-hypothesis, the assumption was tested that the benefits of dynamic over static visualizations should be more pronounced for learners with weaker spatial abilities. Moreover, the cognitive processes associated in learning with the different instructional formats were examined by means of think-aloud protocols (Ericsson & Simon, 1993). The think-aloud protocols in turn were coded according to the taxonomy of learning strategies by Weinstein and Mayer (1986). Thereby claims were examined that can be derived from the cognitive functions that text, dynamic and static visualizations may play in learning. More precisely, it was assumed that adding visualizations to text would offload a learner's working memory, would reduce uncertainty about the content and would lead to a more thorough processing of the content. With regard to the comparison of dynamic and static visualizations, it was assumed that dynamic visualizations would further lessen the demands on working memory and reduce the uncertainty about the content.

These research questions were examined in a one-factorial design with three conditions (text-only, text with dynamic visualizations and text with static visualizations) as independent variable, and spatial abilities as a continuous factor. The think-aloud protocols assessed during the learning phase, the use of the learning environment, learning outcome measures (verbal factual knowledge, pictorial recall and transfer tasks), as well as subjective ratings of extraneous and germane cognitive load served as dependent variables.

Results revealed that visualizations in general were helpful to achieve a deeper understanding (as measured by transfer tasks), as well as to develop a better pictorial mental model (as measured by pictorial recall tasks) of the content as opposed to only text, which is well in line with the outcome-oriented view of the CTML (Mayer, 2001, 2005a, 2009). This observed multimedia effect can be regarded as a crucial precondition for investigating more differentiated research questions concerning the effectiveness of dynamic and static visualizations, since it is ensured that learners have to rely on visualizations in order to come to a comprehensive understanding of the domain at hand. Moreover, through adding visualizations to text, processing demands were reduced (as measured by subjective ratings of cognitive load), which is in line with a functional view of learning with text and visualizations, and specifically with the CLT (e.g., Sweller et al., 1998). Furthermore, as can be reasoned from a functional view, learners were less uncertain about the content (as measured by less negative monitoring statements) and, more importantly, generated more inferences (as measured by the coded learning strategies of the think-aloud protocols).

With respect to dynamic and static visualizations, contrary to the assumption that dynamic visualizations would be more apt than static ones, no differences were observable for any leaning outcome measure between these types of visualizations in Study 1, which will be discussed in more detail in section 7.4. Also, no differences were observable with regard to working memory demands, as measured by the subjective cognitive load ratings. Regarding the role of spatial abilities, results revealed that even though spatial abilities did not moderate learning with dynamic and static visualizations with regard to learning outcomes, learners receiving static visualizations at least played the visualizations more often than learners receiving dynamic visualizations. This might be interpreted as an attempt to compensate for the demands to mentally animate the content, and somewhat corresponds to the ability-as-compensator hypothesis (cf. Mayer & Sims, 1994). Think-aloud protocols revealed that learners with dynamic visualizations made less erroneous statements about the content, and were more confident to have understood the content, as measured by more positive monitoring statements. Since these positive monitoring statements had no relation to any learning outcome measure, this might be interpreted as an illusion of understanding in learning with dynamic visualizations, thereby confirming the results of a study by Lewalter (2003). On the other hand, the analysis of the think-aloud protocols also indicated that this possible illusion of understanding had no negative impact on valuable processing activities (i.e., elaborations), since these were not reduced when learning from dynamic visualizations compared to static visualizations.

To sum up, even though the multimedia effect could be confirmed in Study 1, no differences occurred between dynamic and static visualizations. A reason for the observed instructional equality of these types of visualizations might be that the potentials of dynamic visualizations might not have properly unfold in Study 1. This in turn might be the case, since problems associated with a high degree of visual complexity might have hampered learning with dynamic visualizations. Hence in Study 2 and Study 3, it was investigated whether particularly learning from dynamic visualizations could be optimized by reducing the problems associated with a high degree of visual complexity.

More precisely, in Study 1, the text was presented in written form, mainly because think-aloud protocols were assessed. However, due the assumed high degree of visual complexity in dynamic visualizations, the potential of dynamic visualizations may not properly unfold when learners have to split their attention between written text and visualizations. This inter-representational split of attention can be overcome by using spoken text. In general, according to the modality effect (e.g., Ginns, 2005; Mayer, 2009, Sweller et al., 1998), multimedia instruction can be enhanced by using spoken instead of written text. The prediction that learning from visualizations in general could be facilitated by using spoken text was investigated in Study 2. Furthermore, the assumption was tested whether specifically the benefits of dynamic over static visualizations may become more evident when using spoken instead of written text. This might be the case, because an inter-representational split-attention effect, which is caused by using written text (cf. Schnotz & Lowe, 2008; Sweller et al., 1998) may specifically harm learning with dynamic visualizations, which are assumed to possess a high degree of visual complexity. Moreover, even though in Study 1 no differences were observable between dynamic and static visualizations, it was assumed that in Study 2 dynamic visualizations should be better suited for gaining a deeper understanding of the content. This was supposed to be the case, because several changes between the material in Study 1 and 2 were realized. These changes mainly aimed at making better use of the potentials of dynamic visualizations, such as emphasizing dynamic features in the dynamic visualizations. Accordingly, it was tested whether learning with dynamic visualizations would offload working memory and particularly improve performance on transfer tasks that require a deeper understanding of these dynamic features (cf. Bétrancourt & Tversky, 2000). Furthermore, it was investigated if the benefits of dynamic visualizations would be more pronounced for learners with weaker spatial abilities.

The abovementioned research questions were examined in a 2x2-design with type of visualization (dynamic vs. static) and text modality (written vs. spoken) as independent variables, and spatial abilities as a continuous factor. Dependent variables were learning outcome measures (verbal factual knowledge, pictorial recall, transfer knowledge) and subjective ratings of cognitive load.

Results revealed that adding spoken instead of written text to visualizations fostered learning, and as predicted, specifically with regard to pictorial recall and transfer tasks. Moreover, and in contrast to Study 1, dynamic visualizations were shown to be more apt for conveying a deeper understanding of the content than static visualizations, as measured by transfer tasks. In line with this finding, dynamic visualizations seemed to reduce processing demands, as indicated by subjective ratings of cognitive load. However, contrary to what had been expected, the superiority of dynamic visualizations was not more pronounced when the visualizations were accompanied by spoken instead of written text. There was also no moderating role of spatial abilities in learning with these types of visualizations. Summing up, it occurred that the modifications of the multimedia instruction, and particularly the improvement of the dynamic visualizations to better unfold their potential to convey dynamic features, led to the superiority of dynamic as opposed to static visualizations. Nevertheless, one drawback of dynamic visualizations themselves might still have been apparent (as indicate by the moderate performance on transfer tasks), namely their inherent visual complexity that is, for instance, caused by multiple changes occurring in parallel. To help learners in coping with the visual complexity of dynamic visualizations in particular, but also as a means to improve multimedia instruction in general, it has been suggested to apply cueing (e.g., de Koning et al., 2009). Cueing should thereby facilitate the processes of selection, organization, and integration, and, in connection, should counteract the demands that constitute the high degree of visual complexity in dynamic visualizations. Thus, for Study 3, it was examined if cueing would enhance learning with this kind of multimedia instruction. Thereby, it was expected that cueing would generally lead to better performance in pictorial and transfer tasks. Moreover, it was tested whether the superiority of dynamic over static visualizations for transfer tasks would be more pronounced under cued than under non-cued conditions. Furthermore, to ensure that the superiority of dynamic over static visualizations for transfer tasks found in Study 2 would not only be restricted to one specific presentation format of static visualizations, namely static-sequential visualizations, static-simultaneous visualization conditions were additionally implemented. Finally, it was once again tested if the benefits of dynamic visualizations would particularly hold true for learners with weaker spatial abilities. These assumptions were expected to be mirrored by the respective pattern of cognitive load.

To test these research questions, a 2x3-design was chosen with cueing (with/without) and type of visualizations (dynamic, static-sequential, static-simultaneous) as independent variables, and spatial abilities as a continuous factor. Learning outcome measures (verbal factual knowledge, pictorial recall, transfer knowledge) and subjective ratings of cognitive load served as dependent variables.

Results revealed that learners in the cued conditions developed better pictorial mental models, as indicated by their performance on pictorial recall tasks, and stated that they could better concentrate on the content than learners in the uncued conditions. However, other than expected, cueing neither had an effect on transfer tasks, nor did it decrease ECL. With respect to the type of visualizations, learners receiving dynamic visualizations outperformed learners receiving static visualizations for transfer tasks. Also, dynamic visualizations seemed to reduce processing demands, as indicated by the cognitive load ratings. Thereby, the results of Study 3 mirrored the results of Study 2 with regard to the comparison of dynamic and static visualizations. The presentation format of static visualizations, on the other hand, had no influence on any learning outcome measure, indicating that the presentation format of static visualization played a subordinate role for the understanding of this domain and the dynamic features it comprises. Contrary to the assumption that the superiority of dynamic over static visualizations would be even more pronounced if cueing was implemented, cueing did not moderate learning with dynamic and (different formats of) static visualizations. This may suggest that overall the visual complexity played a subordinate role in learning with the used visualizations. Once again, no moderating role of spatial abilities was observable in learning with dynamic and static visualizations.

In a nutshell, the results of these three studies indicate that first, adding visualizations to text is essential for conveying the abstract Newton's laws of motion by means of a contextualized example, such as the physical principles underlying fish locomotion. Second, particularly dynamic visualizations seem to be better suited than static visualizations for gaining a deeper understanding of the domain given that conditions are met under which dynamic visualizations can unfold their potential. Thereby, it should, for instance, be ensured that the potential of dynamic visualizations to depict dynamic features is exploited. Also, boundary conditions that might hamper learning with dynamic (and static) visualizations, such as inter-representational split-attention or a high visual complexity should be diminished, for instance by using spoken text or by cueing visualizations, even though it should be noted that cueing was solely partly beneficial.

In the following, the research questions and the corresponding results of the current thesis will be discussed in more detail: The benefits of multimedia learning, learning with dynamic and static visualizations, optimizing learning from visualizations, and the moderating role of learner characteristics. Thereafter, some limitations of the current thesis will be discussed, and finally conclusions will be derived.

7.3 The Benefits of Multimedia

As indicated by a vast majority of research (e.g., Anglin et al., 2004; Carney & Levin, 2002; Fletcher & Tobias, 2005; Levie & Lentz, 1982; Levin et al., 1987; Mayer, 2001, 2009), learning with text and visualizations (i.e., multimedia), as opposed to learning with text-only, can be considered as a successful way to enhance learning. There are at least two perspectives to explain why learning with multimedia should be beneficial (cf. Schmidt-Weigand & Scheiter, 2011), namely an outcome-oriented view, as well as a functional view. An outcome-oriented view focuses on the mental representations that are build when learning with multimedia; its most prominent exponent is the CTML (2001, 2005a, 2009). According to the CTML, learning with text and visualizations as opposed to only text should mainly lead to a better pictorial mental model (as measured by pictorial recall tasks) as well as to a better integrated mental model (as measured by transfer tasks), but not necessarily to a better verbal mental model (as measured by verbal factual knowledge tasks). A functional view, on the other hand, focuses on the cognitive processes that are facilitated when learning with text and visualizations. In line with a functional view, learning with text and visualizations as opposed to only text should offload working memory (i.e., decrease extraneous cognitive load), lead to less uncertainty about the content (as reflected by fewer erroneous, fewer negative monitoring and more positive monitoring statements), as well as to support a more elaborate processing of the content (as reflected by more generated inferences, more activated knowledge and an increase in germane cognitive load).

The results of Study 1 are in line with what would be expected from the CTML: Learners receiving text and visualizations outperformed learners receiving only text for pictorial recall and transfer tasks; however, there were no differences for verbal factual knowledge tasks. These results thereby also basically mirror the research on the multimedia effect that differentiates between different knowledge tasks: In a review of his own studies, Mayer (2001) reported higher effect sizes for transfer tasks than for verbal factual knowledge tasks. Similarly, studies that assessed verbal and pictorial factual knowledge tasks found the multimedia effect to be especially pronounced for pictorial tasks, but less accentuated and sometimes even nonexistent for verbal tasks (e.g., Alesandrini & Rigney, 1981; Baker & Dwyer, 2000; Beagles-Roos & Gat, 1983; Joseph & Dwyer, 1984; Szabo et al., 1981; for an overview see Levie & Lentz, 1982).

With respect to the functional view, the results of Study 1 gave only partial support to the derived predictions. Other than expected, there were no differences for the amount of erroneous statements, positive monitoring statements and statements about the activation of knowledge. However, it should be noted that these categories of the coded think-aloud protocols had no substantial relation to learning outcomes. Accordingly, they may not be valid indicators for the cognitive processes suggested by the functional view. As assumed, learning with text and visualizations reduced the processing demands on learners, as indicated by lower subjective ratings of ECL. Moreover, adding visualizations to text led to fewer negative monitoring statements as well as to more generated inferences, both categories that had a substantial relation to learning outcomes. These results confirm findings by Butcher (2006), who also observed more generated inferences for learners receiving text and visualizations. However, the more generated inferences in Study 1 were not reflected by an increase in GCL, which in turn might also be a problem of assessing different types of cognitive load by means of subjective measures (cf. de Jong, 2010), an issue that will be discussed in more detail in section 7.7. It should be noted though that beyond the study of Butcher (2006), the empirical support for the derived assumptions from the functional view is rather sparse, mainly because of a lack of research in this field. For instance, the finding that learners receiving text and visualizations conducted fewer negative monitoring statements can be regarded as a new, first evidence for the claim that adding visualizations to text might lead to less uncertainty regarding the content as compared to text alone.

Hence, to further examine the functions of visualizations in facilitating cognitive processes, further research is needed that applies on-line measures such as think-aloud protocols. Moreover, other online-measures, such as eye-tracking data might additionally contribute to a better understanding of the processes, for instance by relating the time spend on visualizations, or the switches between text and visualizations, with categories of verbal data. Thereby, claims derived from a functional view can be examined in more detail and can enrich an outcome-oriented view (e.g., van Gog, Paas, & van Merriënboer, 2005; see also Scheiter & van Gog, 2009). Likewise extending the methodological repertoire in research on multimedia learning may enrich our understanding of the functions of dynamic and static visualizations and of the cognitive processes they may facilitate.

7.4 The Superiority of Dynamic over Static Visualizations

Concerning the research on the effectiveness of learning with dynamic as opposed to static visualizations, the pattern of results has been rather inconclusive so far: Whereas Tversky et al. (2002) had a rather discouraging view, the meta-analysis of Höffler and Leutner (2007), as well the research overview given in Chapter 2.2.4 of this thesis support an overall advantage of dynamic over static visualizations. However, this is not to be misunderstood as to say that dynamic visualizations are globally always better suited than static visualizations. Rather, it is recommended to take different boundary conditions into account when and why dynamic visualizations should be superior (e.g., Bétrancourt, 2005; Hegarty, 2004; Schnotz & Lowe, 2008). This was done for the three studies of the current thesis. Thereby, it was reasoned that because the domain chosen for this thesis comprises several dynamic features, dynamic visualizations should be better suited than static ones for conveying knowledge in this domain. Thereby, dynamic visualizations were supposed to offload working memory, since resource-intensive processes, specifically spatial and temporal inferences, did not need to be conducted. This superiority was assumed to become true for transfer tasks (cf. Bétrancourt & Tversky, 2000), and was, in line with the ability-as-compensator hypothesis, expected to be more pronounced for learners with weaker spatial abilities.

In Study 1 of this thesis, no differences between dynamic and static visualizations could be observed with regard to any learning outcome measure. On the other hand, for Study 2 and Study 3, as predicted, learners receiving dynamic visualizations performed better on transfer tasks than learners receiving static visualizations. However, several changes in the learning material were realized between Study 1 as opposed to Study 2 and 3: For instance, the complexity of the learning material was reduced, the redundancy of text and visualizations was diminished, and the visualizations changed by adding landscape background to make dynamic features in the dynamic visualizations easier discernable for the learner. Therefore, it is not retraceable which change or combination of changes, respectively, caused these differences between Study 1 compared to Study 2 and 3. It should be noted though that these changes were implemented since they were supposed to generally improve the instructional material, and particularly dynamic visualizations. For this superordinate goal of optimizing learning with dynamic and static visualizations, it was tolerated that results of Study 1 were not directly comparable to Study 2 and Study 3 anymore (see also section 7.7.2).

Irrespective of that, in Study 2 and 3, the benefits of dynamic visualizations became especially evident for transfer tasks, which asked for a deeper understanding of the content, but not for verbal or pictorial factual knowledge tasks. The fact that differences between dynamic and static visualizations solely had an impact on transfer tasks is well in line with theoretical considerations by Bétrancourt and Tversky (2000), who argue that differences between these types of visualizations may mainly become evident in tasks where inferences from a mental model have to be drawn. Thereby, the results stress the importance to distinguish these different learning outcome measures when investigating differences between dynamic and static visualizations. Also, for Study 2 and 3, and corresponding to the results of the transfer tasks, dynamic visualizations seemed to reduce processing demands, as indicated by the subjective cognitive load ratings. This corresponds to the claim that the dynamic visualizations may be beneficial for learners, because they reduce resource-intensive processing demands during learning, namely the need to conduct temporal and spatial inferences (cf. Schnotz & Lowe, 2008).

The major aim of this thesis was to improve multimedia instruction in general, and dynamic visualizations in particular, to convey Newton's laws of motion in the context of the physical principles underlying fish locomotion. Nevertheless, in the current thesis it was also aimed at uncovering boundary conditions at which the benefits of dynamic as opposed to static visualizations might become more or less pronounced. At this, it was argued that due to the assumed higher degree of visual complexity in dynamic as opposed to static visualizations, the benefits of dynamic visualizations would particularly unfold with spoken text. This assumption was tested in Study 2 of the current thesis. However, even though spoken text was beneficial, the benefits of dynamic as opposed to static visualizations were not differently pronounced with regard to text modality. Therefore, it was reasoned that it may be the case that reducing processing demands by using spoken text was not sufficient to deal with the visual complexity of dynamic visualizations per se. Hence, in the following Study 3, the role of cueing – which is supposed to help dealing with visual complexity in (dynamic) visualizations – was examined, and it was argued that the benefits of dynamic over static visualizations would be more pronounced for cued visualizations. Even though cueing helped learners to better recall the information depicted in the visualizations, and dynamic visualizations once again proved to be better suited than static visualizations for achieving a better understanding of the content, no moderating role of cueing in learning with dynamic and static visualizations could be observed.

Taken together, even though the two treatments of spoken text (Study 2) and cueing (Study 3) enhanced learning with both kinds of visualizations (cueing had at least a positive impact on pictorial recall tasks in study 3), both treatments did not accentuate these benefits comparatively stronger in learning with dynamic visualizations. In hindsight, these findings may be interpreted as suggesting that the visual complexity might have not been the major problem in learning with dynamic as opposed to static visualizations. As discussed in Study 3 of this thesis, this might be the case for at least two reasons with regard to the design of the dynamic visualizations used in the current study: First, the content was segmented, that is the visualizations gradually built up from section to section (cf. segmenting principle, Mayer, 2009). Second, the fish's undulatory movement was shown repeatedly in the dynamic visualizations, so that the learners had the chance to see the depicted elements several times. These factors might have already reduced the visual complexity typically observed in many dynamic visualizations, so that it may be argued that the visual complexity has played a subordinate role in processing the current materials. This is not to be misinterpreted in a way that the visual complexity could not be regarded as higher in dynamic as opposed to static visualizations, but solely that it might not have been the major problem.

It should be noted that therefore, with regard to these two factors, the design of Study 2 and Study 3 can be considered as conservative for testing a moderating role of text modality or cueing, respectively, in learning with dynamic and static visualizations. One would still expect to find such a moderating role, when the visual complexity is higher because of a non-segmented or transient display, where a lack of attention is associated with a loss of information.

Overall, it might be concluded that for a dynamic domain like the one at hand, dynamic visualizations might be beneficial for learners to get a deeper understanding of the content – at least if their potential to depict the interrelations of dynamic features is exploited, as might be the case for Study 2 and Study 3 of this thesis. Thereby, with dynamic visualizations a learner does not need to conduct resource-intensive processes, namely temporal and spatial inferences (cf. Schnotz & Lowe, 2008). This is in line with the pattern of results for the item supposed to measure ECL in Study 2 and 3. On the other hand, it was initially assumed that the freed cognitive resources would be dedicated to more valuable processing activities that are associated with an increase in GCL. There might be at least two reasons for why such an effect was missing: First, it may be that an increase in GCL does not happen automatically when learning with dynamic visualizations. Rather, it may be necessary to prompt the learners to actively process the visualizations (e.g., de Koning et al., 2011b; Kombartzky et al., 2010) to devote the freed resources of working memory to GCL when learning with dynamic visualizations. Second, as abovementioned, the lack of an effect for GCL may also be traced back to problems in measuring cognitive load by means of subjective ratings (cf. de Jong, 2010; Schnotz & Kürschner, 2007; see section 7.7.1).

It should be noted that the effect size concerning the superiority of dynamic over static visualizations – although these comparisons were significant in both studies – dropped from $\eta^2 p$ = .12 in Study 2 to $\eta^2 p$ = .03 in Study 3. This might be traced back to the fact that the static visualizations, particularly the static-sequential visualizations, were improved from Study 2 to

Study 3. The improvement in Study 3 was realized by synchronizing the corresponding key frame with the auditory text in case the text referred to a specific state depicted in a key frame, irrespective of whether cueing was implemented or not. However, in Study 2, the static-sequential visualizations were not synchronized with the text to avoid a confounding of synchronization and text modality. Similarly, no synchronization was realized for the dynamic visualizations in Study 2, and for the uncued dynamic visualizations in Study 3, but only for the cued dynamic visualizations in Study 3, which were paused in important states, if the auditory text referred to these states[27]. Therefore, it may be argued that overall, if anything, in Study 3 the static visualizations were favored as opposed to the dynamic visualizations with regard to the synchronicity with the spoken text. Nevertheless, also for Study 3 dynamic visualizations proved to be better suited than static visualizations; however, the effect size was less pronounced compared to Study 2.

Based on these results, it can be argued that dynamic visualizations are better suited than static visualizations for conveying the physical principles underlying fish locomotion, since this domain comprises dynamic features that are crucial for achieving a deeper understanding of the content. However, it would be desirable to be able to generalize these findings with instructional material in another domain that also possess dynamic features like changes in velocity. This was, for instance, recently done in a study by Kühl et al. (2010) for the domain of Kepler's second law. However, once again, this is not to be misunderstood that as long as dynamic features like changes in velocity are crucial for understanding a content, dynamic visualizations are always better than static visualizations, since also then boundary conditions still have to be taken into account. For instance, in the case of the study by Kühl et al. (2010), the superiority of dynamic over static visualizations to convey a deeper understanding of Kepler's second law was only observable if text and (dynamic) visualizations were not redundant.

Hence, and to emphasize this point, the results of the current studies are not to be misinterpreted in a way that dynamic visualizations are globally better than static visualizations (cf. Bétrancourt, 2005; Hegarty, 2004; Scheiter & Gerjets, 2010; Tversky et al., 2002). Rather, dynamic visualizations might be better suited than static visualizations under certain boundary conditions, for instance, if the content comprises crucial visuo-spatial changes and/or dynamic features that are depicted in the dynamic visualizations, and if the respective learning outcome measure asks for the understanding of these properties (cf. Bétrancourt & Tversky, 2000). However, if there is no clear reasoning for why dynamic visualizations should be beneficial, one

[27] It was decided not to pause the dynamic visualizations in the uncued condition, since some authors argue that pausing dynamic visualizations can already be considered as a form of cueing (cf. Schnotz & Lowe, 2008)

consequently should not expect this to be the case! For instance, when considering classic multimedia material used by Mayer and his co-workers (e.g., Mayer et al., 2005), such as “How lightning works”, it is questionable if there are good arguments to expect dynamic visualizations to be beneficial. This is questionable, because on the one hand, the dynamic visualizations neither depict changes that are assumed to be hard to mentally animate (e.g., a cloud rising up in the sky), nor do they depict crucial dynamic features, such as changes in velocity. Moreover, besides hardly possessing an obvious benefit, the dynamic visualizations for this content may be even harmful, since they are transient, thereby possibly imposing unnecessary processing demands on a learner.

Summing up, even though dynamic visualizations might possess enormous potential to particularly convey a domain with dynamic features (e.g., changes in velocity), still then boundary conditions have to be taken into account, such as the learning objective to be achieved by presenting dynamic visualizations to learners.

7.5 Optimizing Learning From Visualizations

To optimize learning from visualizations in general, and dynamic visualizations in particular, two design characteristics were implemented that aimed at reducing inter-representational split-attention by using spoken text (Study 2), and at dealing with the visual complexity of (dynamic) visualizations by means of cueing (Study 3). However, since both design factors did not pronounce the superiority of dynamic over static visualizations differently, in the following the influence of these design factors on learning with the instructional material will be summarized for the two types of visualizations.

As explicated in Chapter 4.1, according to the modality effect in multimedia learning, using spoken instead of written text should lead to better learning outcomes (cf. Ginns, 2005; Low & Sweller, 2005; Mayer, 2009; Sweller et al., 1998). More precisely, and as can be derived from the CTML (Mayer, 2001, 2005a, 2009), the modality effect should result in a better pictorial mental model (as measured by pictorial recall tasks), as well as in a better integrated mental model (as measured by transfer tasks), but not necessarily result in a better verbal mental model. With regard to the demands on working memory, and in line with the CLT (e.g., Sweller et al., 1998), using spoken text should decrease ECL and increase GCL.

As explicated in Chapter 4.2, and in line with the CTML (Mayer, 2001, 2005a, 2009), cueing visualizations in a way that the processes of selection, organization and integration are supported, should also lead to a better pictorial mental model, as well as in a better integrated mental model, but not necessarily in a better verbal mental model. Concerning CLT (e.g., Sweller et al., 1998), cueing should thereby lead to a decrease in ECL and increase in GCL.

The results of Study 2 as well as Study 3 both revealed no differences for verbal factual knowledge tasks, which are supposed to be an indicator for the verbal mental model. However, as expected, accompanying visualizations by spoken text (Study 2), led to better performances for pictorial recall tasks, which is in line with the few studies that also assessed pictorial tasks (e.g., Craig et al., 2002; Mayer & Moreno, 1998; Moreno & Mayer, 1999; Rummer et al., 2011; Schmidt-Weigand et al., 2010; Schüler et al., 2011). Moreover, as predicted, accompanying visualizations by spoken text led to better performance for transfer tasks, which is in line with the results of the meta-analysis by Ginns (2005). With respect to the demands on working memory, other than expected, using spoken instead of written text neither decreased ECL nor increased GCL as measured by subjective cognitive load ratings. This problem of measuring cognitive load by subjective ratings will be discussed in more detail in section 7.7.1.

Concerning cueing, as expected, cueing visualizations (Study 3) led to better performance for pictorial recall tasks, which is in line with the assumptions derived from the CTML as well as with the few studies that examined the influence of cueing on pictorial tasks (e.g., Beck, 1987; Boucheix & Guignard, 2005; Ozcelik et al., 2010; Van Meter et al., 2010). However, other than expected, there were no statistical significant differences between cued and uncued conditions for transfer tasks (Study 3). Even though this latter result is often found in learning with (dynamic) visualizations (cf. de Koning et al., 2009), it is nevertheless somewhat discouraging, because the cues in Study 3 were also designed to support learners in relating information from text and visualizations. This in turn was supposed to result in a better integrated mental model, and finally in better performance on transfer tasks. One possible explanation for the missing effect of cueing with respect to transfer tasks might be that whereas cueing basically helped in relating information from text and visualizations, it may not have supported learners to develop a more elaborated integrated mental model from which inferences concerning interrelations of dynamic features could be drawn. However, such inferences were required to successfully accomplish the transfer tasks in Study 3. Another possible explanation might be that the cueing methods that aimed at supporting learners to relate text and visualizations were not sufficient and should be further enhanced. For instance, it might be beneficial to temporarily add labels to the visualizations to make the relationship of text and visualizations more salient.[28] With respect to the influence of cueing on cognitive load, contrary to the assumptions, cueing did not decrease ECL. However, cueing lead to a marginal increase in GCL. Irrespective of the latter result, doubts

[28] It should be noted though that this might on the other hand increase the visual complexity of the visualizations, which was the main reason to not implement labels in Study 3.

concerning the valid assessment of different types of cognitive load by means of subjective ratings have recently been raised so that the interpretation of these items should be treated cautiously (cf. de Jong, 2010). The problem of assessing cognitive load will be discussed in more detail in section 7.7.1.

Overall, a focus in the current thesis was to optimize learning with visualizations in general, and dynamic visualizations in particular. Even though learning with visualizations could be optimized by means of using spoken text and, at least partly, by cueing, this improvement was not more pronounced for learning with dynamic visualizations. However, there may be more important variables – which were not explicitly investigated in the current thesis – that may optimize learning with (dynamic) visualizations, which will be discussed in the following.

It is occasionally recommended to improve learning with dynamic visualizations through implementing interactivity (e.g., Hasler, Kersten, & Sweller, 2007; Schwan & Riempp, 2004; Tabbers & de Koeijer, 2010), particularly if the dynamic visualizations are transient. However, even for transient visualizations, the research on interactivity is not conclusive (cf. Boucheix, 2008): While in some studies a positive effect of interactivity was found (e.g., Tabbers & de Koeijer, 2010; Wang et al., 2011), in other studies interactivity in learning with dynamic visualizations had no influence (e.g., Boucheix & Guignard, 2005; Boucheix & Schneider, 2009; Exp. 2; Kriz & Hegarty, 2007), or even a negative effect (Bétrancourt & Réalini, 2005). Particularly, for the chosen domain, namely the physical principles underlying fish locomotion, implementing interactivity might have been harmful. This is due to the fact that if interactivity was implemented in a way that learners would be able to rewind the visualizations (e.g., by means of a slider), this would likely result in a misconception: When playing the visualizations backwards, one would not only see a "wrong" undulatory movement of a fish, but more importantly, the concept of a reaction force would be depicted erroneous, because in this case, the movement of a body segment and the movement of the reaction force would point in the same direction instead of opposite directions! Hence, the only reasonable interactivity would be to implement a play, pause, and replay button (i.e., self-pacing). However, as abovementioned, since the dynamic visualizations were not transient, but were displaying the fish's movement repeatedly, the need for self-pacing was considered to play a subordinate role. Nevertheless, it cannot be ruled out that self-pacing might have led to a different pattern of results in Study 2 and Study 3. For instance, the modality effect, which was observed in Study 2 (Study 2 was system-paced), might have not been observed under self-paced conditions, since self-pacing might diminish the modality effect (cf. Ginns, 2005; Tabbers, 2002). Also, the influence of cueing – which was at least found for pictorial tasks – might have been diminished if self-pacing had been applied, since learners could have taken all the time they needed to view the visualizations and extract the information, so that, for instance, the attention-guiding function of cueing might have been less important. On the other hand, it should be noted that it probably would take more time to learn under self-paced than system-paced conditions. Hence, even though self-pacing may make cueing unnecessary, since it allows achieving the same *effectiveness* this would not mean that self-pacing would also be equally *efficient* compared to the abovementioned design characteristics. However, while these assumptions are plausible, they need to be investigated in further studies, thereby possibly also taking into account the modality of the text and the visualizations' transience, since this may further influence the effect.

Moreover, pacing might not only affect the modality effect and cueing, but it might also influence the relative effectiveness of learning with dynamic and static visualizations. Whereas in Study 2 and 3 dynamic visualizations were better suited for the domain at hand than static visualizations, this was not the case for Study 1. Several things changed between Study 1 as opposed to Study 2 and 3: the complexity of the domain in general, the design of the visualization and what they depicted, the redundancy between text and visualizations, the application of think-aloud protocols, but also if the learning environment was self-paced (Study 1) or system-paced (Study 2 and 3). Even though it is not retraceable what exactly led to the different pattern of results, it cannot be completely ruled out that solely the change from self-pacing to system-pacing accounts for this finding (see also Höffler, 2007; Exp. 2 & 3). This might be the case because learners with static visualizations may have tried to compensate for the demands of mentally animating the changes when learning with this kind of visualizations. This notion gets further indirect support from the fact that in Study 1 particularly learners with weaker spatial abilities receiving static visualizations played the visualizations more often, thereby possibly trying to compensate the demands of mental animation. Hence, the kind of pacing may moderate the effectiveness in learning with dynamic and static visualizations in that benefits of dynamic over static visualizations become less pronounced once self-pacing is implemented. In that case, static visualizations may be as effective as dynamic visualizations, but less efficient, since learners with static visualizations might take longer time for learning. The relative effectiveness and efficiency of static and dynamic visualizations under conditions of self-pacing as opposed to system-pacing should be addressed in future research.

A major goal of the current thesis was to optimize learning with multimedia instruction, at which the focus was set on *improving the design* of the multimedia material in general, and the design for learning with dynamic visualizations in particular. However, this is only one side of the coin. The other side of the coin may be to *improve processing activities*, that is, to encourage learners to more adequately and thoroughly process the content. Several methods have been suggested to do so, such as self-explanations (e.g., de Koning et al., 2011b; Gerjets, Scheiter, &

Catrambone, 2006; Renkl, 2002; for an overview see Chi, 2005), prompts (e.g., Berthold, Nückles, & Renkl, 2007; Glogger, Schwonke, Holzäpfel, Nückles, & Renkl, 2009), or the teaching of learning strategies (e.g., Kombartzky et al., 2010; Schlag & Plötzner, 2010; Selcuk, Sahin, & Acikgöz, 2011). By teaching learners how to apply learning strategies, a learner's understanding can be improved, as recently shown by Kombartzky et al. (2010) as well as Schlag and Plötzner (2010). The effectiveness of such an approach may be accentuated differently for learning with dynamic and static visualizations. For instance, learning with dynamic visualizations might lead to an illusion of understanding as indicated by the results of Study 1 of this thesis, and by Lewalter (2003) – even though it did not lead to a shallower processing. Nevertheless, it might be the case that when learners are aware of such an illusion and are taught how to process dynamic visualizations in a way that they can devote their freed resources to a more thorough processing, they might benefit more from such a learning strategy than their counterparts learning with static visualizations. On the other hand, it can also be construed that when learners would be taught how to (or be prompted to) conduct spatial and temporal inferences, one may assume that such a strategy might be more beneficial for (novice) learners receiving static visualizations, since these learners are supposed to struggle in conducting such inferences in particular. Summing up, providing adequate learning strategies possesses enormous potential to improve a learner's understanding. Whether and how such learning strategies play a moderating role in learning with static and dynamic visualizations, is assumed to depend on characteristics of the visualizations, for instance, what they depict, as well as on the properties of the learning strategies, such as which kind of cognitive processes they might stimulate. Research in this field is still at its beginning and it is exciting what further studies in this area will reveal.

7.6 The Moderating Role of Learner Characteristics

In the current thesis it was not only investigated how to optimize learning with dynamic and static visualizations, but also the influence of learner characteristics, specifically spatial abilities, in learning with these types of visualizations was examined. Concerning the learner characteristic spatial ability, according to the ability-as-compensator hypothesis (cf. Mayer & Sims, 1994), particularly learners with weaker spatial abilities should profit from learning with dynamic visualizations (cf. Höffler, 2010). This was assumed to hold true for Studies 1, 2, and 3. However, for all three studies, spatial abilities did not moderate learning with dynamic and static visualizations for any learning outcome measure, independently if the static visualizations were shown sequentially or simultaneously (Study 3). Only for Study 1, and in accordance with the ability-as-compensator hypothesis, spatial abilities moderated the frequency of playing dynamic and static visualizations, respectively. In addition, for all three studies, a main effect of spatial abilities on learning outcomes could be observed, in that stronger spatial abilities were associated with higher learning outcomes, irrespective of the type of visualization. This main effect of spatial abilities is in line with the current status of research (cf. Höffler, 2010), and emphasizes the importance of assessing spatial abilities in learning with visualizations.

It should be noted though that even there are good theoretical reasons, meta-analytic empirical evidence (Höffler, 2010), as well as evidence by a few studies (e.g., Boucheix & Schneider, 2009; Höffler, 2007) to expect that spatial abilities should moderate learning with dynamic and static visualizations, overall this effect is nevertheless rarely found in single studies (cf. Hegarty & Kriz, 2008). At this, the three studies of this thesis are in line with the majority of published studies concerning this topic, indicating that the moderating role of spatial abilities in learning with dynamic and static visualizations may be a fragile effect, only occurring under certain circumstances.

A possible explanation for the lack of a moderating role of spatial abilities on learning with dynamic and static visualizations with regard to learning outcomes for the studies in the current thesis (even though it should be noted that in Study 1 spatial abilities moderated the frequency of playing the visualizations), might be traced back to the applied test to measure spatial abilities, namely the MRT (Vandenberg & Kuse, 1978). The MRT is supposed to mainly measure a person's ability to mentally rotate objects, but not necessarily the ability to infer dynamic features and interrelations. However, the dynamic visualizations in the current thesis did not only depict visuo-spatial changes, but also dynamic features. Hence, to compensate for these differences between dynamic and static visualizations, learners with static visualizations were not only required to mentally animate the spatial changes, but also to infer dynamic features with static visualizations to achieve a similar level as their counterparts receiving dynamic visualizations (cf. Schnotz & Lowe, 2008). Since the MRT might not match best with the task demands posed by learning with static visualizations, it hence might not have been apt to reveal a moderating influence for the dynamic and static visualizations used in these studies. Similarly, Höffler (2007; see also Stebner, 2009) stressed the importance of an appropriate test to observe a moderating effect in learning with dynamic and static visualizations. For instance, Höffler (2007; Exp. 2) observed a moderating role of spatial abilities in learning with dynamic and static visualizations solely for a test that belonged to the factor visualization (VZ), but not for a test that belonged to the factor spatial relation (SR), and concluded that the test belonging to VZ matched best the demands of mental animation for his used instructional material. For the used instructional material of the current thesis, a well-suited test might not only measure the ability to reason about spatial changes, but also the ability to reason about dynamic features. Hence, it may not be necessary to reject the “ability-as-compensator-hypotheses” per se; rather spatial abilities may simply not be the ability to best match the task demands imposed by learning with static and dynamic visualizations for this kind of instructional material. A further line of future research with regard to the influence of spatial abilities may pay closer attention to the fact that the dynamic information has to be perceived when learning with dynamic visualizations. Hence, one might additionally consider tests that aim at measuring this ability. According to D’Olivera (2004), such a test might be a dynamic spatial ability test, since “dynamic spatial abilities refers to the ability to deal with moving elements and relative motion” (p. 20; see also Hegarty & Waller, 2005).

Another crucial learning prerequisite, which might have consequences on the effectiveness in learning with dynamic and static visualizations, and which was investigated in several prior studies, is a learner’s expertise (e.g., ChanLin, 1998, 2001; Kalyuga, 2008, Schnotz & Rasch, 2005; Yarden & Yarden, 2010; Zhu & Grabowski, 2006). In the experiments of the current thesis, basically novice learners were investigated, but not experts, who in this case would most likely be people with a background in Physics. This means that a moderating role of expertise, or prior knowledge respectively, was not explicitly investigated (cf. prior knowledge principle [or expertise reversal effect, respectively], Kalyuga, 2005). Nevertheless, as a learners’ expertise can be regarded as an important factor in learning with dynamic and static visualizations, a test for assessing the prerequisite knowledge of learners was applied in all three studies to control for this factor.

Concerning a learner’s expertise, there are mostly two lines of reasoning with respect to how it influences learning with dynamic and static visualizations. One line of reasoning argues that a higher expertise is necessary to make sense of dynamic visualizations, as they would otherwise be overwhelming (e.g., Kalyuga, 2008; Schnotz & Rasch, 2005). Some studies support this view: Low prior knowledge students learned better with static than with dynamic visualizations, whereas there were no differences between the two types of visualizations for high prior knowledge students (e.g., ChanLin, 2001; Kalyuga, 2008). Importantly, for these studies there was no clear reasoning why dynamic visualizations should be beneficial. Rather, in these cases dynamic visualizations could even be regarded as harmful, since they did not depict crucial visuo-spatial changes or dynamic features, and, were moreover transient, thereby imposing unnecessarily processing demands onto learners. Accordingly, high prior knowledge may in these studies simply have served to compensate for the negative and unnecessary demands imposed by dynamic visualizations that low prior knowledge learners suffered from.

The other line of reasoning corresponds to the ability-as-compensator hypothesis. Thereby, it is argued that learners with a high expertise might learn equally well with static compared to dynamic visualizations, because they might be able to conduct valuable processing activities, such as inferring dynamic features, which in turn might be conducive to building an adequate mental model. Novice learners, on the other hand, are not supposed to be able conduct such processes (e.g., Boucheix & Guignard, 2005; Lowe, 1996), but nevertheless are able to perceive the dynamic features depicted in dynamic visualizations. Accordingly, one may assume that in this case the benefits of dynamic compared to static visualizations might become especially evident for novice learners. This pattern of results was found in a study by Yarden and Yarden (2010): Dynamic visualizations were more apt than static visualizations for learners with low prior knowledge, while there were no differences for learners with high prior knowledge. Note that the ability-as-compensator hypothesis with regard to a learner’s expertise might be reasonable if dynamic visualizations depict dynamic features, like for the multimedia instruction in the studies of this thesis. Accordingly, for the novice learners investigated in this thesis, dynamic visualizations were more apt than static visualizations (Study 2 and 3). Hence, one may wonder whether dynamic visualizations would still be better suited than static visualizations if learners possessed high prerequisite knowledge with regard to their ability to easily infer dynamic features, so that these processes would not be regarded as resource-demanding anymore. Moreover, one may speculate whether for these learners the need to actively process static visualizations (e.g., by conducting spatial and temporal inferences, instead of perceiving all the relevant processes in the dynamic visualizations), would even be more beneficial. Similar effects have been observed in text comprehension research, where high prior knowledge learners benefitted more strongly from incoherent rather than coherent instructional text (McNamara, Kintsch, Songer, & Kintsch, 1996). The underlying explanation for these effects is that high prior knowledge learners are better able to conduct inferences to overcome the information gaps in the text and that these inferences yield a deeper comprehension than what could be achieved from a more coherent text. Similarly, high prior knowledge learners may profit more strongly from (incoherent) static visualizations, which promote their inference activities. Among yielding deeper comprehension of the content, these inferences may also pay off in the long run, for instance, in delayed testing by yielding a more durable mental representation of the content (cf. Palmiter & Elkerton, 1993). Hence, for future research, one might test if this observed superiority of dynamic visualizations to depict dynamic features might also account for learners possessing a high expertise. Also, future research might consider further learner characteristics. For instance, even though spatial abilities as well as a learner’s expertise are the two predominantly examined learner characteristics in learning with dynamic and static visualizations, recently Höffler et al. (2010) examined the influence of cognitive style (cf. Massa & Mayer, 2006; Mayer & Massa, 2003). Höffler et al. (2010) found a moderating effect in that high developed visualizers learning with static visualizations outperformed their counterparts learning with dynamic visualizations, whereas there were no differences between the types of visualizations for low developed visualizers. Summing up, to be able to identify under which conditions which kind of learner would benefit most from which type of visualizations, future research should also consider different learner characteristics in learning with dynamic and static visualizations (cf. Hegarty, 2004; Tversky et al., 2002).

7.7 Limitations of the Current Studies

In the following, limitations of the current studies will be discussed, particularly the assessment of cognitive load and the changes in the instructional material between the three studies.

7.7.1 Cognitive Load Measurement

In the current thesis, in all three studies, based on CLT, the demands on working memory associated in learning with this kind of multimedia instruction were examined by means of subjective cognitive load ratings (e.g., Sweller et al., 1998). It should be noted though that there are problems associated with the measurement of cognitive load. Several suggestions have been made how to measure cognitive load and its components, thereby taking into account objective (cf. Brünken, Seufert, & Paas, 2010; Paas, Tuovinen, Tabbers, & van Gerven, 2003) as well as subjective measurements, the latter being predominant ones when measuring cognitive load in educational research (cf. Paas et al., 2003; van Gog & Paas, 2008). The attractiveness of subjective measures might be traced back to the fact that they are relatively easy to implement and, furthermore, do not interfere with the learning process itself. Probably the most often used measurement for measuring cognitive load is the mental-effort item developed by Paas (1992). It should be noted that instead of measuring cognitive load during learning, the mental effort item was originally developed to measure the effort for solving problems or knowledge tests respectively (cf. van Gog & Paas, 2008)[29]. However, when applied for measuring cognitive load, one critical drawback of the mental-effort item is that it does not distinguish between the three load types explicitly, which makes the interpretation of this item problematic (cf. de Jong, 2010). Therefore, it is desirable to have subjective rating measures available that can distinguish

[29] In this context, it is also often used to calculate the instructional efficiency of a given multimedia instruction.

between, and map consistently to, the different load types (cf. Brünken et al., 2010; de Jong, 2010; Gerjets, Scheiter, & Cierniak, 2009). At this, several attempts have been made to formulate items that may be apt to differentiate between the different load types (e.g., Cierniak et al., 2009; Corbalan, Kester, & van Merriënboer, 2008; Scheiter et al., 2006; Gerjets et al., 2009), with partly inconclusive results (cf. de Jong, 2010). Hence, also in the current thesis, it was tried to use subjective measures that are able to distinguish between the different load types.

In Study 1, two items were assessed to measure ECL ("How difficult was it for you to understand the contents?") and GCL ("How much effort did you invest in order to understand the content?"), respectively. However, since particularly the item supposed to measure GCL seemed to be insufficient, for Study 2 and Study 3 two items were used from a study by Cierniak et al. (2009), which originally could be mapped successfully to ECL ("How difficult was it for you to learn with the given material?") and GCL ("How much did you concentrate during learning?"), respectively. The results for these two items, which were assessed in Study 2 and 3 of this thesis, however, were only partly in line with what would be expected from CLT. While ratings on the ECL-item were consistently lower in the dynamic than in the static visualization conditions, thereby reflecting the results for learning outcomes in a theory-consistent way, the pattern of results concerning the modality effect and cueing was not in line with CLT. The GCL-item was basically in line with CLT for cueing, but not in line with the CLT with regard to the modality effect, making it arguable if this item indeed measured GCL. There were also no differences for the GCL-item with respect to type of visualization. However, there are at least two explanations for this latter result: First, it may be that even if dynamic visualizations reduce ECL, they do not necessarily lead to an increase in GCL. Second, it might be that dynamic visualizations also lead to an increase in GCL, but that no effect is observable for the GCL-item due to problems in measuring cognitive load (cf. de Jong, 2010).

It should be noted though that measuring different types of cognitive load by means of subjective ratings may be fragile. At this, occasionally doubts have been raised, if learners are generally able to distinguish between the load types by means of subjective evaluation and/or if these subjective ratings, which are assessed after a learning phase really reflect the processing demands during learning (e.g., de Jong, 2010; Schnotz & Kürschner, 2007). Therefore, the items should be treated with caution, and should not be assessed uncritically. It should also be noted that even though several more or less successful attempts have been made so far, no measure that might be mapped to different load types won ultimate recognition yet (cf. de Jong, 2010). Thus, although the assessed items in the study might not perfectly match ECL and GCL, respectively, they nevertheless can be considered as one of the most suited subjective ratings in

CLT research so far. On the other hand, this also indicates that research on measuring components of cognitive load is in need of improvement (cf. de Jong, 2010).

7.7.2 Comparability Across Studies

In the current thesis, learning with dynamic and static visualizations was examined across three consecutive studies. Thereby, conditions that were used in one experiment were also used in the following experiment. Nevertheless, it must be admitted that these studies are not directly comparable, because the instructional material changed from study to study in other potentially important variables. This incomparability was accepted for the superordinate goal of optimizing the instructional material that was newly developed for the current thesis. In Study 1, the instructional material left room for improvement in several ways: For instance, to have a fair comparison between the text-only and the text and visualization conditions, the text was extensive, thereby leading to a high redundancy between text and visualizations. This redundancy, however, might be even harmful for learning from multimedia (Schmidt-Weigand & Scheiter, 2011; Sweller, 2005b). Moreover, the potential of the dynamic visualizations to depict dynamic features and their interrelation to other variables was not completely exploited for the multimedia instruction in Study 1, which might account for the instructional equality of dynamic and static visualizations in this study. Hence, to further improve the multimedia instruction, for Study 2, several changes were realized: For instance, to counteract the abovementioned drawbacks the redundancy of text and visualizations was reduced, and the potential of dynamic visualizations to depict dynamic features and interrelations was accentuated, thereby possibly mostly enhancing learning with dynamic visualizations. Moreover, other changes occurred between Study 1 and Study 2, such as the reduction of the complexity of the domain itself, changes in the knowledge test, the pacing of the instructional material or the applying of think-aloud protocols. Due to these changes, it cannot be decided what caused dynamic visualizations to be superior to static visualizations in Study 2, but not in Study 1[30]. Similarly, also Study 2 and 3 were not directly comparable. First, some minor changes were realized with regard to the domain itself, as well as to the related knowledge test. Moreover, and probably more importantly, also the synchronization between spoken text and particularly static-sequential visualizations was improved from Study 2 to Study 3. Thereby, particularly the static-sequential visualization

[30] Note that it would have needed to use a 2x2x2x2x2x2-design with 64 conditions to unravel this discrepancy, with redundancy, design of visualizations, complexity of domain, changes in the knowledge test, pacing, and applying of think-aloud protocols as independent variables. This was considered as too extensive for the scope of this thesis, especially when keeping in mind that the superordinate goal was to improve the instructional material.

condition was favored as compared to the dynamic visualization condition, which in turn may have led to a drop in the effect size for the superiority of dynamic over static visualizations. To sum up, even though in general it might be preferable to conduct studies that are directly comparable, in the current thesis an incomparability was taken into account for the superordinate goal of optimizing learning from dynamic and static visualizations.

7.8 Conclusion

All in all, the current thesis aimed at finding design factors that would optimize learning with text and visualizations in general, and dynamic visualizations in particular as compared to static visualizations. Study 1 proved that adding visualizations to text improved comprehension. Study 2 showed that reducing inter-representational split-attention by using spoken text was beneficial for learning with dynamic as well as with static visualizations. Study 3 gave at least partly support for the assumption that cueing would improve learning with dynamic as well as static visualizations. Moreover, as the results for Study 2 and Study 3 indicated, dynamic as compared to static visualizations supported learners to gain a deeper understanding of the used domain, thereby stressing the importance of using different knowledge tasks, such as transfer tasks. Even though spatial abilities were an important learner characteristic, the superiority of dynamic visualizations was not more pronounced for learners with weaker spatial abilities.

As abovementioned, future studies should examine if the observed superiority of dynamic over static visualizations can be generalized to other domains that possess similar properties as the domain of the physical principles underlying fish locomotion (e.g., Kepler's second law). Moreover, it would be interesting to know whether the results would also hold true if interactivity or self-pacing, respectively, would be implemented, and whether this design factor would even further improve the multimedia instruction. Hereby, it is argued that such predictions seem to be solely reasonable when task characteristics of the used multimedia instruction as well as learner characteristics are taken into account. Furthermore, for ongoing studies, one might not only focus on design characteristics of the multimedia instruction, but also on how to engage learners in processing the instructional material more thoroughly. These issues might help us in achieving a better understanding on when and how to support which kind of learners, so that these learners can benefit most when dealing with a given subject.

8 References

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

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183-198.

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*, 669-681.

Alesandrini, K. L., & Rigney, J. (1981). Pictorial presentation and review strategies in science learning. *Journal of Research in Science Teaching*, *18*, 465-474.

Amadieu, F., Mariné, C., & Laimay, C. (2011). The attention-guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior, 27*, 36-40.

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

Ardac D., & Akaygun S. (2005). Using static and dynamic visuals to represent chemical change at molecular level. *International Journal of Science Education, 27*, 1269-1298.

Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior*, *25*, 354-359.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (vol. 2, pp. 89-195). New York: Academic Press.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94,* 416-427.

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior*, *25*, 348-353.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21*, 695-700.

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 134-146). New York: Cambridge University Press.

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

Baek, Y. K., & Layne, B. H. (1988). Color, graphics, and animation in a computer assisted learning tutorial lesson. *Journal of Computer-Based Instruction, 15,* 131-135.

Baker, R., & Dwyer, F. (2000). A meta-analytic assessment of the effect of visualized instruction. *International Journal of Instructional Media, 27*, 417-426.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612-637.

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.

Beagles-Roos J., & Gat, I. (1983). Specific impact of radio and television on children's story comprehension. *Journal of Educational Psychology, 75,* 128-137.

Beck, C. R. (1984). Visual cueing strategies: Pictorial, textual, and combinational effects. *Educational Communication & Technology Journal, 32*, 207-216.

Beck, C. R. (1985). The effects of pictorial and textual cues on learning intentional and incidental information. *International Journal of Instructional Media, 12*, 299-311.

Beck, C. R. (1987). Pictorial cueing strategies for encoding and retrieving information. *International Journal of Instructional Media, 14*, 332-346.

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17,* 564-577.

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology, 101*, 70-87.

Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287-296). New York: Cambridge University Press.

Bétrancourt, M., & Réalini, N. (2005, June). *Faut-il vraiment laisser le déroulement d'une animation sous le contrôle de l'apprenant?* Paper presented at the 11th Journées d'Etude sur le Traitement Cognitif des Systèmes d'Information Complexes (JETCSIC), Nice, France.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Le travail humain 63*, 311-330.

Blake, T. (1977). Motion in instructional media: Some subject-display mode interactions. *Perceptual and Motor Skills, 44*, 975-985.

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Auflage). Heidelberg: Springer Verlag.

Boucheix, J.-M. (2008).Young learners' control of technical animations. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (pp. 208-234). New York: Cambridge University Press.

Boucheix, J.-M., & Guignard, H. (2005). What animated illustrations can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education, 4*, 369-388.

Boucheix, J.-M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction, 20*, 123-135.

Boucheix, J.-M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction, 19,* 112-127.

Bransford, J., & Johnson, M. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.

Brünken, R., & Leutner, D. (2001). Aufmerksamkeitsverteilung oder Aufmerksamkeitsfokussierung? Empirische Ergebnisse zur ‘‘Split-Attention-Hypothese’’ beim Lernen mit Multimedia. [Split attention or focusing of attention? Empirical results on the split-attention hypothesis in multimedia learning]. *Unterrichtswissenschaft, 29,* 357-366.

Brünken, R., Seufert, T., & Paas, F. (2010). Measuring cognitive load. In: J. Plass, R. Moreno & R. Brünken (Eds.), *Cognitive Load: Theory & Application* (pp.181-203). New York: Cambridge University Press.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology, 49,* 109-119.

Brünken, R., Steinbacher, S., Schnotz, W., & Leutner, D. (2001). Mentale Modelle und Effekte der Präsentations- und Abrufkodalität beim Lernen mit Multimedia [Mental models and the effects of presentation and retrieval mode in multimedia learning]. *Zeitschrift für Pädagogische Psychologie, 15,* 16-27.

Butcher, K. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*, 182-197.

Camtasia (Version 3.0) [computer software]. East Lansing, MI: TechSmith.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students’ learning from text. *Educational Psychology Review, 14,* 5-26.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Catrambone, R., & Seay, A. F. (2002). Using animation to help students learn computer algorithms. *Human Factors, 44,* 495-511.

ChanLin, L.-J. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology, 25,* 166-175.

ChanLin, L.-J. (2000). Attributes of animation for learning scientific knowledge. *Journal of Instructional Psychology, 27,* 228-238.

ChanLin, L.-J. (2001). Formats and prior knowledge on learning in a computer-based lesson*.* *Journal of Computer-Assisted Learning, 17,* 409-419.

Chi, M. T. H. (2005). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 271-286). New York: Cambridge University Press.

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.

Cohen, C. A., & Hegarty, M. (2007). Individual differences in use of an external visualization while performing an internal visualization task. *Applied Cognitive Psychology, 21*, 701−711.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2008). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *Contemporary Educational Psychology, 33*, 733-756.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94,* 428-434.

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology, 35*, 59-74.

D'Oliveira, T. C. (2004). Dynamic spatial ability: An exploratory analysis and a confirmatory study. *International Journal of Aviation Psychology, 14,* 19-38.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*, 105-134.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21*, 731-746.

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 Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010a). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction, 20*, 111-122.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010b). Learning by generating vs. receiving instructional explanations: Two approaches to enhance attention cueing in animations. *Computers & Education, 55*, 681-691.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011a). Attention cueing in an instructional animation: The role of presentation speed. *Computers in Human Behavior, 27*, 41-45.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011b). Improved effectiveness of cueing by self-explanations when learning from a complex animation. *Applied Cognitive Psychology, 25*, 183-194.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (revised edition)*. Cambridge, MA: MIT Press.

Erkens, G. (2005). *Multiple episode protocol analysis (MEPA). Version 4.10.* Utrecht University, The Netherlands.

Fischer, S. (2008). Temporal Manipulations in Instructional Animation Design: Is Attention Guiding Thought? (Band 10 der Reihe Wissensprozesse und digitale Medien). Berlin: Logos.

Fischer, S., Lowe, R. K., & Schwan, S. (2008). Effects of presentation speed of a dynamic visualization on the understanding of a mechanical system. *Applied Cognitive Psychology, 22*, 1126-1141.

Fischer, S., & Schwan, S. (2010). Comprehending animations. Effects of spatial cueing versus temporal scaling. *Learning and Instruction, 20*, 465-475.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 117-134). New York, NY: Cambridge University Press.

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text-segmentation, picture labeling, and spatial proximity. *Learning and Instruction, 20*, 216-224.

Folker, S., Ritter, H., & Sichelschmidt, L. (2005). Processing and integrating multimodal material: The influence of color-coding. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the cognitive science society* (pp. 690-695). Mahwah, NJ: Erlbaum.

Franik, R., & Klose, U. (2007). *Duden – Naturwissenschaft und Technik.* Berlin: Duden Paetec Schulbuchverlag.

Gerjets, P., Imhof, B., Kühl, T., Pfeiffer, V., Scheiter, K., & Gemballa, S. (2010). Using static and dynamic visualizations to support the comprehension of complex dynamic phenomena in the Natural Sciences. In L. Verschaffel, E. de Corte, T. de Jong, & J. Elen (Eds.), *Use of* *external representations in reasoning and problem solving: Analysis and improvement* (pp. 153-168). London: Routledge.

Gerjets, P., & Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load: Evidence from hypertext-based instruction. *Educational Psychologist, 38*, 33-41.

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction, 16*, 104-121.

Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of Cognitive Load Theory: A research agenda based on the Structuralist View of Theories. *Educational Psychology Review, 21*, 43-54.

Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. S. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior, 25*, 360-370.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313-331.

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2009). Activation of learning strategies in writing learning journals: The specificity of prompts matters. *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology, 23*, 95-104.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69-110.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*, 462-466.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*, 64-74.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*, 713-729.

Hays, T. A. (1996). Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *Journal of Educational Computing Research, 14,* 139-155.

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

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14,* 343-351.

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., & Kozhevnikov, M. (1999). Spatial abilities, working memory and mechanical reasoning. In J. Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design* (pp. 221-241). Sydney, Australia: University of Sydney.

Hegarty, M., & Kriz, S. (2008). Effects of knowledge and spatial ability on learning from animation. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (pp. 3-29). New York: Cambridge University Press.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21,* 325-360.

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

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121-169). New York: Cambridge University Press.

Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science, 30*, 581-592.

Höffler, T. N. (2003). Animation, Simulation oder Standbilder? Zur Wirkung interaktiver Animationen auf das Verständnis komplexer biologischer Prozesse [Animation, simulation or static pictures? Concerning the effectiveness of interactive animations upon the comprehension of complex biological processes] (Unpublished diploma thesis). University of Kiel, Germany.

Höffler, T. N. (2007). Lernen mit dynamischen Visualisierungen: Metaanalyse und experimentelle Untersuchungen zu einem naturwissenschaftlichen Lerninhalt (Electronical dissertation). Universität Duisburg-Essen, Germany.

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

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17,* 722-738.

Höffler, T. N., Prechtl, H., & Nerdel, C. (2010). The influence of visual cognitive style when learning from instructional animations and static pictures. *Learning and Individual Differences, 20*, 479-483.

Huk, T. (2006). Who benefits from learning with 3d models? The case of spatial ability. *Journal of Computer-Assisted Learning, 22*, 392-404.

Huk, T., Steinke, M., & Floto, C. (2010). The educational value of visual cues and 3D-representational format in a computer animation under restricted and realistic conditions. *Instructional Science, 38*, 455-469.

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*, 172-176.

Imhof, B., Scheiter, K., & Gerjets, P. (2009). Realism in dynamic, static-sequential, and static-simultaneous visualizations during knowledge acquisition on locomotion patterns. In N. A. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2962-2967). Austin, TX: Cognitive Science Society.

Imhof, B., Scheiter, K., Gerjets, P., & Edelmann, J. (2010). The role of dynamic visualizations and spatial layout of static visualizations for learning how to classify locomotion patterns. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2039-2044). Austin, TX: Cognitive Science Society.

Iskander, W., & Curtis, S. (2005). Use of colour and interactive animation in learning 3D vectors. *Journal of Computers in Mathematics and Science Teaching*, *24*, 149-156.

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*, 135-145.

Janelle, C. M., Champenoy, J. D., Coombes, S. A., & Mousseau, M. B. (2003). Mechanisms of attentional cueing during observational learning to facilitate motor skill acquisition. *Journal of Sport Sciences*, *21,* 825-838.

Jarodzka, H. (2011). In the eye of an expert: Conveying perceptual skills in biological and medical domains via eye movement modeling examples (Unpublished dissertation). Eberhard-Karls-Universität Tübingen, Germany.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*, 146-154.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators on dual sensory mode instruction. *Educational Psychology, 17*, 329-343.

Joseph, J. H., & Dwyer, F. M. (1984). The effects of prior knowledge, presentation mode, and visual realism on student achievement. *Journal of Experimental Education, 52*, 110-121.

Kalyuga, S. (2005). Prior knowledge in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 325-338). New York: Cambridge University Press.

Kalyuga, S. (2008). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior, 24,* 852-861.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13,* 351-371.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92,* 126-136.

Keller, T., Gerjets, P., Scheiter, K., & Garsoffky, B. (2006). Information visualizations as tools for knowledge acquisition: The impact of dimensionality and color coding. *Computers in Human Behavior, 22*, 43-65.

Kim, S., Yoon, M., Whang, S.-M., Tversky, B., & Morrison, J. B. (2007). The effect of animation on comprehension and interest. *Journal of Computer-Assisted Learning, 23*, 260-270.

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

Koroghlanian, C., & Klein, J. D. (2004). The effect of audio and animation in multimedia instruction. *Journal of Educational Multimedia and Hypermedia, 13,* 23-46.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*, 911-930.

Kühl, T., Scheiter, K., & Gerjets, P. (2010, August). *Can text compensate differences in types of graphics?* Poster presented at the meeting of the EARLI SIG2 Text and Graphics Comprehension, Tübingen, Germany.

Kühl, T., Scheiter, K., & Gerjets, P. (2011). *The impact of cueing in learning with dynamic and static visualizations.* Manuscript in preparation.

Kühl, T., Scheiter, K., Gerjets, P., & Edelmann, J. (2011). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior, 27*, 29-35.

Kühl, T., Scheiter, K., Gerjets, P., & Gemballa, S. (2011). Can differences in learning strategies explain the benefits of learning from static and dynamic visualizations? *Computers & Education, 56*, 176-187.

Lai, S.-L. (2000). Increasing associative learning of abstract concepts through audiovisual redundancy. *Journal of Educational Computing Research, 23,* 275-289.

Large, A., Beheshti, J., Breuleux, A., & Renaud, A. (1996). Effect of animation in enhancing descriptive and procedural texts in a multimedia learning environment. *Journal of the American Society for Information Science, 47,* 437-448.

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

Leahy, W., Chandler, P., & Sweller, J. (2002). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology, 17*, 401-418.

Levie, W., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology Journal, 30,* 195-232.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration* (vol. 1, pp. 51-85). New York: Springer.

Levin, J. R., & Divine-Hawkins, P. (1974). Visual imagery as a prose-learning process. *Journal of Reading Behavior, 6,* 23-30.

Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction, 13*, 177-189.

Lin, H., Chen, T., & Dwyer, F. M. (2006). Effects of static visuals and computer-generated animations in facilitating immediate and delayed achievement in the EFL classroom. *Foreign Language Annals, 39*, 203-219.

Lin, H., & Dwyer, F. M., (2010). The effect of static and animated visualization: a perspective of instructional effectiveness and efficiency. *Educational Technology Research and Development, 58,* 155-174.

Lohman, D. F., Pellegrino, J. W., Alderton, D. L., & Regian, J. W. (1987). Dimensions and components of individual differences in spatial abilities. In S. H. Irvine & S. N. Newstead (Eds.), *Intelligence and cognition: Contemporary frames of reference* (pp. 253-312). Dordrecht, Netherlands: Nijhoff.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology, 75*, 402-412.

Lorch, R. F., & Lorch, E. P. (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology, 88*, 38-48.

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147-158). New York: Cambridge University Press.

Lowe, R. K. (1996). Background knowledge and the construction of a situational representation of a diagram. *European Journal of Psychology of Education, 11,* 377-397.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *14*, 225-244.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13,* 247-262.

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14,* 257-274.

Lowe, R. K., Schnotz, W., & Rasch, T. (2011). Aligning affordances of graphics with learning task requirements. *Applied Cognitive Psychology, 25*, n/a. doi:10.1002/acp.1712

Marbach-Ad, G., Rotbain, Y., & Stavy, R. (2008). Using computer animation and illustration activities to improve high school students' achievement in molecular genetics. *Journal of Research in Science Teaching, 45*, 273-292.

Massa, L. J., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences, 16,* 321-335.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*, 377-389.

Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology, 99*, 640-652.

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

Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge: University Press.

Mayer, R. E. (2005b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1-16). Cambridge: University Press.

Mayer, R. E. (2005c). Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 183-200). Cambridge: University Press.

Mayer, R. E. (2009). *Multimedia learning (second edition)*. New York: Cambridge University Press.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: what works in the design of agent-based microworlds? *Journal of Educational Psychology, 95,* 806-813.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*, 256-265.

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95,* 833-841.

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, 87,* 312-320.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load. *Educational Psychologist, 38*, 43-52.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dualcoding theory of multimedia learning. *Journal of Educational Psychology, 86,* 389-401.

McNamara, D., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14,* 1-43.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* New York: Elsevier.

Meyer, K., Rasch, T., & Schnotz, W. (2010). Effects of animation's speed of presentation on perceptual processing and learning. *Learning and Instruction*, *20*, 136-145.

Moore, P. J., & Scevak, J. J. (1997). Learning from texts and visual aids: A developmental perspective. *Journal of Research in Reading, 20*, 205-223.

Moreno, R. (2007). Optimizing learning from animations by minimizing cognitive load: Cognitive and affective consequences of signaling and segmentation methods. *Applied Cognitive Psychology, 21*, 765-781.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*, 358-368.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94,* 598-610.

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

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319-334.

Münzer, S., Seufert, T., & Brünken, R. (2009). Learning from multimedia presentations: Facilitation function of animations and spatial abilities. *Learning and Individual Differences, 19,* 481-485.

Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies, 57,* 279-315.

Nerdel, C. (2003). Die Wirkung von Animation und Simulation auf das Verständnis von stoffwechselphysiologischen Prozessen [The influence of animation and simulation on the comprehension of physiological metabolism] (Electronical dissertation). University of Kiel, Germany.

Nicholls, C., & Merkel, S. (1996). The effect of computer animation on students' understanding of microbiology. *Journal of Research on Computing in Education, 28,* 359-371.

O'Neil, H. F., Mayer, R. E., Herl, H., Niemi, C., Olin, K., & Thurman, R. A. (2000). Instructional strategies for virtual environments. In H. F. O'Neil, & D. H. Andrews (Eds.), *Aircraft training: Methods, technologies, and assessment* (pp. 105-130). Mahwah, NJ: Erlbaum.

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26,* 110-117.

Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding affects multimedia learning. *Computers & Education, 53*, 445-453.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429-434.

Paas, F., Tuovinen, J. E., Tabbers, H. K., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71.

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

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255-287.

Palmiter, S., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction, 8*, 193-216.

Park, O.-C., & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays – a review. *Instructional Science, 21,* 427-449.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction, 3,* 227-238.

Pfeiffer, V. D. I., Gemballa, S., Jarodzka, H., Scheiter, K., & Gerjets, P. (2009). Situated learning in the mobile age: Mobile devices on a field trip to the sea. *Research in Learning Technology, 17*, 187-199.

Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educational effective animations and simulations. *Journal of Computing in Higher Education, 21*, 31-61.

Rebetez, C., Bétrancourt, M., Sangin, M., & Dillenbourg, P. (2010). Learning from animation enabled by collaboration. *Instructional Science, 38*, 471-485.

Renkl, A. (2002). Learning from worked-out examples: Instructional explanations supplement self-explanations. *Learning and Instruction, 12,* 529-556.

Resnick, L.B. (1987). Learning in school and out. *Educational Researcher, 16,* 13-20.

Rieber, L. P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research, 5,* 431-444.

Rieber, L. P. (1990). Using computer animated graphics with science instruction with children. *Journal of Educational Psychology, 82,* 135-140.

Rieber, L. P. (1991). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology, 83,* 318-328.

Rieber, L. P., Boyce, M. J., & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction, 17,* 46-52.

Rieber, L. P., & Kini, A. S. (1991). Theoretical foundations of instructional applications of computer-generated animated visuals . *Journal of Computer-Based Instruction, 18*, 83-88.

Rigney, J. W., & Lutz, K. A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitude. *Journal of Educational Psychology, 68,* 305-311.

Rummer, R., Schweppe, J., Fürstenberg, A., Scheiter, K., & Zindler, A. (in press). The perceptual basis of the modality effect in multimedia learning. *Journal of Experimental Psychology: Applied*.

Rummer, R., Schweppe, J., Fürstenberg, A., Seufert, T., & Brünken, R. (2010). Working memory interference during processing texts and pictures: Implications for the explanation of the modality effect. *Applied Cognitive Psychology, 24*, 164-176.

Rummer, R., Schweppe, J., Scheiter, K., & Gerjets, P. (2008). Lernen mit Multimedia: Die kognitiven Grundlagen des Modalitätseffekts. *Psychologische Rundschau, 59*, 98-107.

Russell, J., & Hollander, S. (1975). A biology attitude scale. *The American Biology Teacher, 37*, 270-273.

Salomon, G. (1979). *Interaction of media, cognition, and learning: An exploration of how symbolic forms cultivate mental skills and affect knowledge acquisition.* San Francisco: Jossey-Bass.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76,* 647-658.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185-213.

Scheiter, K. (2009). *Theoretical and empirical foundations of theories of multimedia learning: A critical reconsideration* (Unpublished professorial dissertation). Knowledge Media Research Center, Tübingen, Germany.

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 – 6th International Conference, Diagrams 2010* (LNAI 6170, pp. 264-270). Heidelberg: Springer.

Scheiter, K., & Gerjets, P. (2010). Cognitive and socio-motivational aspects in learning with animations: There is more to it than 'do they aid learning or not'. *Instructional Science, 38*, 435-440.

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Visualizing mathematical solution procedures. *Computers in Human Behavior, 22*, 9-26.

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.

Scheiter, K., Wiebe, E., & Holsanova, J. (2008). Theoretical and instructional aspects of learning with visualizations. In R. Zheng (Ed.), *Cognitive effects of multimedia learning* (pp. 67-88). Hershey, PA: IGI Global.

Schlag, S., & Plötzner, R. (2010). Supporting learning from illustrated text: Conceptualizing and evaluating a learning strategy. *Instructional Science*. doi:10.1007/s11251-010-9160-3

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

Schmidt-Weigand, F., & Scheiter, K. (2011). The role of spatial descriptions in learning from multimedia. *Computers in Human Behavior, 27*, 22-28.

Schnotz, W. (2002). 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: Cambridge University Press.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19*, 469-508.

Schnotz, W., & Lowe, R. K. (2003). External and internal representations in multimedia learning. *Learning and Instruction, 13,* 117-123.

Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (pp. 304-356). New York: Cambridge University Press.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology: Research and Development, 53*, 47-58.

Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2011). *Explaining the modality effect in multimedia learning: Is it due to high visuo-spatial load in working memory or a lack of temporal contiguity?* Manuscript submitted for publication.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14,* 293-305.

Schwartz, D. L. (1995). Reasoning about the referent of a picture versus reasoning about the picture as a referent: An effect of visual realism. *Memory & Cognition, 23*, 709-722.

Selcuk, G. S., Sahin, M., & Acikgöz, K. (2011). The effects of learning strategy instruction on achievement, attitude, and achievement motivation in a physics course. *Research in Science Education, 41*, 39-62.

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

Seufert, T., Jänen, I., & Brünken R. (2007). The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence formation. *Computers in Human Behavior, 23*, 1055-1071.

She H.-C., & Chen Y.-Z. (2009). The impact of multimedia effect on science learning: Evidence from eye movements. *Computers & Education, 53,* 1297-1307.

Spangenberg, R. W. (1973). The motion variable in procedural learning. *AV Communication Review, 21,* 419-436.

Spotts, J., & Dwyer, F. M. (1996). The effect of computer-generated animation on student achievement of different types of educational objectives. *International Journal of Instructional Media, 23,* 365-375.

Stebner, F., Lebens, M., Wirth, J., & Opfermann, M. (2009, March). *Learning from animations and static pictures: The impact of spatial ability and cognitive load.* Paper presented at the Third International Cognitive Load Conference, Heerlen, The Netherlands.

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

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

Sweller, J. (2005b).The redundancy principle in multimedia learning multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159-168). New York: University Press.

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

Szabo, M., Dwyer, F., & DeMelo, H. (1981). Visual testing – visual literacy's second dimension. *Educational Communication and Technology Journal, 29*, 177-187.

Szabo, M., & Poohkay, B. (1996). An experimental study of animation, mathematics achievement, and attitude toward computer-assisted instruction. *Journal of Research on Computing in Education, 28,* 390-402.

Taasoobshirazi, G., & Carr, M. (2008). A review and critique of context-based physics instruction and assessment. *Educational Research Review, 3*, 155-167.

Tabbers, H. K. (2002). *The modality of text in multimedia instructions: Refining the design guidelines* (Unpublished doctoral dissertation, Open University of the Netherlands, Heerlen). Retrieved from http://www.ou.nl/Docs/Expertise/OTEC/Publicaties/huib%20tabbers/doctoral%20dissertation%20Huib%20Tabbers%20-%20web%20version.pdf

Tabbers, H. K., & de Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science, 38*, 441-453.

Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2000). *Multimedia instructions and cognitive load theory: split-attention and modality effects.* Paper presented at the National Convention of the Association for Educational Communications and Technology, Long Beach, CA.

Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2001). The modality effect in multimedia instructions. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 1024-1029). Mahwah, NJ: Lawrence Erlbaum Associates.

Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology, 74,* 71-81.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3,* 257-287.

Tunuguntla R., Rodriguez O., Ruiz J. G., Qadri S. S., Mintzer M. J., Van Zuilen M. H., & Roos B. A. (2008). Computer-based animations and static graphics as medical student aids in learning home safety assessment: A randomised controlled trial. *Medical Teacher*, *30,* 815-817.

Tversky, B., Bauer-Morrison, J., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57,* 247-262.

Tversky, B., Heiser, J., Lozano, S., MacKenzie, R., & Morrison, J. (2008). Enriching animations. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (pp. 263-285). New York: Cambridge University Press.

van Gerven, P. W. M. (2002). *Efficient complex skills training into old age: Exploring the benefits of cognitive load theory* (Unpublished doctoral dissertation, Open University of the Netherlands, Heerlen). Retrieved from http://www.personeel.unimaas.nl/p.vangerven/Index_files/Pascal%20van%20Gerven%20-%20Dissertation.pdf van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., & Schmidt, H. G. (2006). Modality and variability as factors in training the elderly. *Applied Cognitive Psychology, 20,* 311-320

van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 1-11.

van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror-neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review, 21*, 21-30.

van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology, 19*, 205-221.

van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*, 237-244.

van Meter, P., Gu, W., Pastore, R., & Cook, M. (2010, August). *Effects of metacognitive instructions and color coding on learnign from multiple representations in biology.* Paper presented at the meeting of the EARLI SIG2 Text and Graphics Comprehension, Tübingen, Germany.

van Oostendorp, H., & Beijersbergen, M. J. (2007, August). *Animated diagrams: Their effect on understanding, confidence and mental effort.* Paper presented at the 12th Biennial meeting of the European Association of Research on Learning and Instruction, Budapest, Hungary.

van Oostendorp, H., Beijersbergen, M. J., & Solimani, S. (2008). Conditions for learning from animations. In P. A. Kirschner, F. Prins, V. Jonker, & G. Kanselaar (Eds.), *International Perspectives in the Learning Sciences: Cre8ing a learning world. Proceedings of the Eighth International Conference for the Learning Sciences – ICLS 2008* (vol. 2, pp. 438-445-397). Utrecht: The Netherlands: International Society of the Learning Sciences, Inc. (ISLS).

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes.* London: Academic Press.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental Rotations, a group test of three-dimensional spatial visualization. *Perceptual Motor Skills, 47*, 599-604.

Waltner, C., Rachel, A., & Wiesner, H. (2006). Computersimulationen von Fortbewegungsarten bei Fischen. In V. Nordmeier & A. Oberländer (Eds.), *Didaktik der Physik – Kassel 2006*. Berlin: Lehmanns Media.

Waltner, C., Wiesner, H., & Rachel, A. (2007). Physics in context – a means to encourage student interests in physics. *Physics Education, 42*, 502-507.

Wang, P.-Y., Vaughn, B.K., & Liu, M. (2011). The impact of interactivity on novices' learning of introductory statistics. *Computers & Education, 56*, 300-311.

Watson, G., Butterfield, J., Curran, R., & Craig, C. (2010). Do dynamic work instructions provide an advantage over static instructions in small scale assembly tasks? *Learning and Instruction, 20*, 84-93.

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

Wells, R. G., van Mondfrans, A. P., Postlethwait, S. N., & Butler, D. C. (1973). Effectiveness of three visual media and two study formats in teaching concepts involving time, space, and motion. *AV Communication Review, 21*, 233-241.

Wender, K. F., & Mühlböck, J.-S. (2003). Animated diagrams in teaching statistics. *Behavior Research Methods, Instruments, & Computers, 35,* 255-258.

Whitelegg, E., & Edwards, C. (2002). Beyond the laboratory: Learning physics using real-life contexts. In Behrendt, H., Dahncke, H., Duit, R., Gräber, W., Komorec, M., Kross, A., & Reiska, P. (Eds.), *Research in science education: Past, present, and future* (pp. 337-342). Dorderecht: Kluwer.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., & Sweller, J. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior, 25*, 339-347.

Wright, P., Milroy, R., & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education, 14,* 203-224.

Yang, E. M., Andre, T., & Greenbowe, T. Y. (2003). Spatial ability and the impact of visualization/animation on learning electrochemistry. *International Journal of Science Education, 25,* 329-349.

Yarden, H., & Yarden, A. (2010). Learning using dynamic and static visualizations: Students' comprehension, prior knowledge and conceptual status of biotechnological method. *Research in Science Education, 40*, 375-402.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Sciences, 18*, 87-122.

Zhu, L., & Grabowski, B. L. (2006). Web-based animation or static graphics: is the extra cost of animation worth it? *Journal of Educational Multimedia and Hypermedia, 15*, 329-347.

Zusammenfassung

In der vorliegenden Arbeit wurde in einer Reihe von 3 Experimenten untersucht in wie fern das Verständnis naturwissenschaftlicher Zusammenhänge allgemein durch den Einsatz von multimedialem Lernmaterial, und im Besonderen durch den Einsatz von dynamischen Visualisierungen (z.B. Videos oder Animationen), optimiert werden kann. Hierfür wurde exemplarisch die Lerndomäne der physikalischen Prinzipien, die einer undulatorischen Fischbewegung zu Grunde liegen, herausgegriffen. Diese Domäne spiegelt Probleme wider, mit denen Lernende in den Naturwissenschaften häufig konfrontiert sind, nämlich zu Verstehen wie eine Veränderung in einer Variablen zu Veränderungen in einer anderen Variablen führen kann (z.B. wie sich die Geschwindigkeitsänderung der Schwanzflosse eines Fisches auf andere Variablen, wie die entstehenden Reaktionskräfte oder die Schwimmgeschwindigkeit auswirken kann).

In Studie 1 wurde untersucht, ob das Hinzufügen von Visualisierungen zu Text generell zu einem besseren Verständnis für die ausgewählte Domäne führen würde. Darüber hinaus wurde getestet, ob – wie aus den Eigenschaften dieser Domäne abgeleitet – dynamische Visualisierungen zu einem tieferen Verständnis führen als statische Visualisierungen. Hierbei wurde zudem überprüft, ob sich mögliche Vorteile dynamischer Visualisierungen gegenüber statischen Visualisierungen im Besonderen für Lernende mit einem geringen räumlichen Vorstellungsvermögen zeigen. Ferner wurden die kognitiven Prozesse beim Lernen mit Text, dynamischen sowie statischen Visualisierungen mittels Laut-Denken-Protokollen untersucht (vgl. Ericsson & Simon, 1993).

Hierfür wurde in Studie 1 ein einfaktorielles Design mit drei Bedingungen (nur Text, Text + dynamische Visualisierungen, Text + statische Visualisierungen) realisiert. Als abhängige Variablen fungierten die erhobenen und kategorisierten Laut-Denken-Protokolle, die Abspielhäufigkeit der Visualisierungen, der Lernerfolg sowie die subjektiv eingeschätzte kognitive Belastung. Es zeigte sich, im Einklang mit der kognitiven Theorie multimedialen Lernens (z.B. Mayer, 2009), dass das Hinzufügen von Visualisierungen zu Text sowohl zu einer besseren Leistung für bildhafte Aufgaben, als auch zu einem tieferen Verständnis (gemessen durch Transferaufgaben) führte. Zudem führte dies zu einer geringer eingeschätzten kognitiven Belastung. Die Auswertung der Laut-Denken-Protokolle zeigte, dass Visualisierungen zu mehr Inferenzen anregten und zu weniger Äußerungen über Verständnisschwierigkeiten führten. Ferner produzierten Lernende mit dynamischen Visualisierungen weniger fehlerhafte Äußerungen und gaben häufiger an die Inhalte verstanden zu haben. Da letzteres jedoch nicht der Fall war, kann dies als eine „Verstehensillusion“ angesehen werden (vgl. Lewalter, 2003). Entgegen der Annahmen gab es jedoch keine Unterschiede zwischen dynamischen und statischen Visualisierungen hinsichtlich des Lernerfolgs und der eingeschätzten kognitiven Belastung.

Ein Grund für die instruktionale Gleichwertigkeit von statischen und dynamischen Visualisierungen in Studie 1 könnte unter anderem daran gelegen haben, dass sich das Potential dynamischer Visualisierungen nicht vollständig entfalten konnte, da Probleme, die mit der vergleichsweisen hohen visuellen Komplexität dynamischer Visualisierungen einhergehen, in Studie 1 gegenwärtig waren. Daher wurde in Studie 2 und 3 untersucht, ob durch die Reduktion der Probleme, die mit einer hohen visuellen Komplexität verbunden sind, das Lernen mit dem benutzten multimedialem Instruktionsmaterial im Allgemeinen, und das Lernen mit dynamischen Visualisierungen im Besonderen, optimiert werden könnte.

In Studie 1 wurde aufgrund der Erhebung von Laut-Denken-Protokollen geschriebener Text dargeboten. Wegen der angenommenen visuellen Komplexität von dynamischen Visualisierungen ist es jedoch möglich, dass sich das Potential dynamischer Visualisierungen nicht entfalten kann, wenn Lernende ihre Aufmerksamkeit zwischen Text und Visualisierungen wechseln müssen. Diese Aufmerksamkeitsteilung zwischen geschriebenem Text und Visualisierung kann mittels gesprochenen Texts beseitigt werden. Gemäß dem Modalitätseffekt (z.B. Sweller et al., 1998) kann allgemein das Lernen mit Multimedia durch Verwendung von gesprochenem statt geschriebenem Text verbessert werden. Der Modalitätseffekt wurde in Studie 2 untersucht. Darüber hinaus wurde untersucht, ob der Vorteil dynamischer Visualisierungen gegenüber statischen Visualisierungen stärker zum Vorschein kommt, wenn gesprochener statt geschriebener Text verwendet wird. Diese Annahme lag darin begründet, dass die Verarbeitung dynamischer Visualisierungen, die aufgrund ihrer visuellen Komplexität viel visuelle Aufmerksamkeit benötigen, durch geschriebenen Text beeinträchtigt wird, so dass ihr Potential durch den Lernenden nicht voll ausgeschöpft werden kann. Ferner wurde angenommen, dass obwohl in Studie 1 keine Unterschiede zwischen dynamischen und statischen Visualisierungen beobachtbar waren, in Studie 2 dynamische Visualisierungen prinzipiell besser wären als statische Visualisierungen um ein tieferes Verständnis der Inhaltsdomäne - gemessen durch Transferaufgaben - zu erlangen (cf. Bétrancourt & Tversky, 2000). Diese Annahme lag darin begründet, da verschiedene Änderungen am Lernmaterial zwischen Studie 1 und 2 vorgenommen wurden, die im Wesentlichen darauf abzielten das Potential dynamischer Visualisierungen besser nutzbar zu machen. Beispielsweise wurde unter anderem das Potential dynamischer Visualisierungen Veränderungen in der Geschwindigkeit darzustellen, und den Einfluss den dies wiederum auf andere Variablen hat, in der ersten Studie nicht ausgereizt. Dies, und noch weitere Veränderungen am Lernmaterial, wurde jedoch für Studie 2 realisiert. Es wurde zudem erwartet, dass das Lernen mit dynamischen Visualisierungen die kognitive Belastung beim Lernenden reduziert. Schließlich wurde getestet, ob der Vorteil dynamischer Visualisierungen stärker für Lernende mit geringerem räumlichem Vorstellungsvermögen ausgeprägt sein würde.

Diese Forschungsfragen wurden mittels eines 2x2-Designs mit Visualisierungsart (dynamisch vs. statisch) und Textmodalität (geschrieben vs. gesprochen) untersucht. Die Ergebnisse zeigten, dass, im Einklang mit dem Modalitätseffekt, gesprochener Text zu einer besseren Leistung speziell für bildhafte Aufgaben und Transferaufgaben führte. Zudem führte das Lernen mit dynamischen Visualisierungen zu einer besseren Leistung bei Transferaufgaben als das Lernen mit statischen Visualisierungen. Im Einklang mit diesem Ergebnis schätzten Lernende die kognitive Belastung beim Lernen mit dynamischen Visualisierungen als geringer ein. Allerdings war, entgegen der ursprünglichen Annahme, der Vorteil von dynamischen gegenüber statischen Visualisierungen nicht stärker für gesprochenen im Vergleich zu geschriebenem Text ausgeprägt. Auch moderierte das räumliche Vorstellungsvermögen nicht das Lernen mit dynamischen und statischen Visualisierungen. Auch wenn dynamische Visualisierungen nun zu einem tieferen Verständnis führten als statische Visualisierungen, so blieb ein Nachteil dynamischer Visualisierungen bestehen, nämlich ihre inhärente visuelle Komplexität.

Um das Lernen allgemein zu verbessern und aber auch speziell der visuellen Komplexität entgegen zu wirken, wird, speziell in letzter Zeit, vorgeschlagen Cueing-Methoden zu verwenden (z.B. de Koning et al., 2009). Daher wurde in Studie 3 untersucht, ob Cueing zum einen das Lernen mit den gegebenen Lernmaterialien verbessern könnte. Zudem wurde angenommen, dass die Überlegenheit von dynamischen gegenüber statischen Visualisierungen stärker unter Cueing-Bedingungen als unter Bedingungen ohne Cueing zum Vorschein kommen würde. Darüber hinaus wurde, um zu gewährleisten, dass der in Studie 2 gefundene Vorteil dynamischer gegenüber statischen Visualisierungen nicht nur auf eine spezielle Form statischer Visualisierungen, nämlich statisch-sequentieller Visualisierungen zurückzuführen ist, zusätzlich noch statisch-simultane Visualisierungen implementiert. Zudem wurde wieder getestet, ob der Vorteil dynamischer Visualisierungen stärker für Lernende mit geringerem räumlichem Vorstellungsvermögen ausgeprägt sein würde.

Diese Annahmen wurden mittels eines 2x3-Designs mit Cueing (ja/nein) und Visualisierungsart (dynamisch, statisch-sequentiell, statisch-simultan) untersucht. Die Ergebnisse zeigten, dass Lernende der Cueing-Bedingungen bessere Leistungen in Bilderaufgaben erzielten als Lernende aus den nicht gecueten Bedingungen. Entgegen der Erwartungen zeigte sich dieser Effekt jedoch nicht für Transferaufgaben. Das Lernen mit dynamischen Visualisierungen führte zu besseren Leistungen in Transferaufgaben als das Lernen mit statischen Visualisierungen, wobei sich keine Unterschiede für die beiden Präsentationsformate statischer Visualisierungen ergaben. Zudem schätzten Lernende der dynamischen Visualisierungsbedingungen die kognitive Belastung als geringer ein als Lernende der statischen Visualisierungsbedingungen. Diese Ergebnisse von Studie 3 bezüglich des Lernens mit dynamischen und statischen Visualisierungen spiegeln dabei die Ergebnisse von Studie 2 wider. Entgegen der ursprünglichen Annahme war die Überlegenheit dynamischer im Vergleich zu statischen Visualisierungen nicht stärker ausgeprägt, wenn die Visualisierungen gecuet waren. Auch wurde das Lernen mit dynamischen und statischen Visualisierungen nicht durch das räumliche Vorstellungsvermögen moderiert.

Knapp zusammengefasst lässt sich festhalten, dass erstens, für die ausgewählte Domäne das Hinzufügen von Visualisierungen zu Text für das Verständnis wesentlich ist. Zweitens, scheinen speziell dynamische Visualisierungen besser geeignet als statische Visualisierungen um zu einem tieferen Verständnis der Domäne zu gelangen – zumindest wenn Bedingungen gegeben sind unter denen sich das Potential dynamischer Visualisierungen entfalten kann. Hierbei sollte beispielsweise gewährleistet sein, dass das Potential dynamischer Visualisierungen um dynamische Eigenschaften darzustellen ausgereizt wird. Auch empfiehlt es sich Bedingungen zu verwenden unter denen (dynamische) Visualisierungen gut verarbeitet werden können, wie beispielsweise durch die Verwendung gesprochenen Textes oder von Cueing-Methoden.

Appendices

Appendix A

Items and Loadings of the Attitudes Towards Biology and Physics Questionnaire on the Assumed Two Factors after Varimax Rotation, Listed as a Function of Factors and Size of Loadings

	Item	Factor 1 (Biology)	Factor 2 (Physics)
1	Ich habe definitiv eine positive Einstellung zur Biologie; sie ist angenehm.	.93	-.16
2	Ich mag Biologie wirklich.	.91	-.13
3	Ich interessiere mich sehr für Biologie.	.88	-.14
4	Bei Biologie fühle ich mich sicher und es regt mich zugleich an.	.80	-.16
5	Ich mag Biologie nicht, und es ängstigt mich es haben zu müssen.	.76	.03
6	Es macht mich nervös, auch nur daran zu denken, ein Experiment in der Biologie durchzuführen.	.66	.29
7	Bei Biologie fühle ich mich unwohl, ruhelos, gereizt und ungeduldig.	.61	-.03
8	Ich fühle mich wohl in Physik und mag sie sehr gerne.	-.05	.92
9	Im Allgemeinen fühle ich mich bezüglich Physik wohl.	-.05	.89
10	Ich gehe an Physik mit einem Gefühl des Zögerns heran.	.08	.87
11	Wenn ich Physik höre, habe ich ein Gefühl der Abneigung.	-.12	.86
12	Physik ist faszinierend und macht Spaß.	-.16	.84
13	Mir hat es in der Schule immer Spaß gemacht Physik zu haben.	-.10	.84
14	Im Fach Physik war/bin ich immer unter großer Anspannung.	.03	.78

Note: Negatively formulated items (items 5, 6, 7, 10, 11, and 14) were recoded

Appendix B

Items and Loadings of the Attitudes Towards Biology and Physics Questionnaire on the Assumed Two Factors after Varimax Rotation, Listed as a Function of Factors and Size of Loadings

	Item	Factor 1 (Biology)	Factor 2 (Physics)
1	Ich mag Biologie wirklich.	.92	-.05
2	Ich habe definitiv eine positive Einstellung zur Biologie; sie ist angenehm.	.92	.06
3	Ich interessiere mich sehr für Biologie.	.88	-.02
4	Es macht mich nervös, auch nur daran zu denken, ein Experiment in der Biologie durchzuführen.	.73	.08
5	Bei Biologie fühle ich mich unwohl, ruhelos, gereizt und ungeduldig.	.73	.02
6	Ich fühle mich wohl in Physik und mag sie sehr gerne.	.01	.90
7	Wenn ich Physik höre, habe ich ein Gefühl der Abneigung.	-.04	.90
8	Physik ist faszinierend und macht Spaß.	.07	.88
9	Im Allgemeinen fühle ich mich bezüglich Physik wohl.	.14	.84
10	Ich gehe an Physik mit einem Gefühl des Zögerns heran.	-.06	.77

Note: Negatively formulated items (items 4, 5, 7, 10) were recoded.