Effects of Task Experience on Attention to Extraneous Information during Multimedia Learning

Effecten van taakervaring op aandacht voor overbodige informatie tijdens het leren met multimedia

Gertjan Rop

Layout by Gertjan Rop
Cover design by Charly Eielts & Gertjan Rop
The orginial artwork by Piet Mondriaan is part of the collection of het Gemeentemuseum Den Haag
Printed by Ridderprint BV | www.ridderprint.nl

ISBN: 978-94-6299-875-9

Effects of Task Experience on Attention to Extraneous Information during Multimedia Learning

Effecten van taakervaring op aandacht voor overbodige informatie tijdens het leren met multimedia

Proefschrift

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam

op gezag van de rector magnificus
Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 5 april 2018 om 13:30 uur

door
Geert Johannes Rop
geboren te Stadskanaal

Erasmus University Rotterdam

Promotiecommissie

Promotor
Prof.dr. T.A.J.M. van Gog

Overige leden
Prof.dr. G.W.C. Paas
Prof.dr. P.C.J. Segers
Dr. H.K. Tabbers

Copromotor
Dr. P.P.J.L. Verkoeijen

Contents

Chapter 1

Introduction

Throughout their lives, people will encounter numerous combinations of words and pictures when gathering information and accumulating new knowledge. For example, in early childhood, children learn the meaning of new words by coupling a spoken word to a picture. In later stages of life, children, adolescents and adults read magazines, newspapers, websites, and study textbooks or e-learning resources that combine text (written or spoken) with pictorial information (static or dynamic). A large body of research exists in educational psychology which tried to identify how multimedia materials (the term used to refer to instructional materials that combine text and pictures) should be designed to optimize learning. This research is inspired by the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2014) and Cognitive Load Theory (CLT; Sweller, Ayres, & Kalyuga, 2011), and has led to the establishment of numerous principles for effective multimedia design.

This dissertation is concerned with the central tenet of two of those principles. The coherence principle (Mayer & Fiorella, 2014), states that people learn more deeply from a multimedia message when unnecessary or irrelevant material is excluded rather than included. The redundancy principle (Kalyuga & Sweller, 2014), suggests that presenting redundant material (e.g. the same information in two different formats) interferes with rather than facilitates learning. In effect, both principles entail that the presentation of extraneous information should be avoided, because it hampers learning compared to instructional materials from which this information has been eliminated. Extraneous information is defined as information that is either irrelevant (i.e., not related to the learning goal) or unnecessary (i.e., related to the learning goal, but not necessary for learning because the information is presented twice or is unnecessarily elaborate). However, eye tracking studies suggest that people can learn to ignore task-irrelevant information and focus more on task-relevant information with relatively little practice (Haider & Frensch, 1999) or explicit instruction (Canham & Hegarty, 2010; Hegarty, Canham, & Fabrikant, 2010). The central question addressed in this dissertation, therefore, is whether extraneous information (either irrelevant or unnecessary) would continue to hamper learning when it is present over a series of tasks, items, or slides, which would give learners the chance to adapt their study strategy. When learners would start to ignore the extraneous information, its negative effect on learning should diminish or disappear.

This question is both theoretically and practically relevant. As for theoretical relevance, investigating this question would provide more insight into task experience (i.e., familiarity with the design of the materials) as a possible boundary condition to the negative effect of extraneous information on learning, and the establishment of boundary conditions is important for describing the limits of generalizability of scientific theories (Busse, Kach, & Wagner, 2016; Whetten, 1989). Moreover, research has long been focused on what multimedia designers can do to aid student learning, and the present studies contribute to a recent, new direction of research focusing on whether students can (learn) to adapt their study strategy, and, thus, self-manage their cognitive load when learning with multimedia (e.g., Agostinho, Tindall-Ford, & Roodenrys, 2013; Gordon, Tindall-Ford, Agostinho, & Paas, 2016; Roodenrys, Agostinho, Roodenrys, & Chander, 2012). In terms of practical relevance, the knowledge gained from addressing this question would be useful for instructional designers. That is, it is very hard for instructional designers to take into account all multimedia principles, because individual learner characteristics may interact with some of the principles (e.g. what is essential information for a novice may be redundant information for a more advanced learner; Kalyuga & Sweller, 2014).

In the remainder of this Chapter, I will first discuss the CTML and CLT in more detail, after which I will give an overview of the literature on the negative effect of extraneous information on learning and the different factors that might influence this effect. Then, I will focus on the effect of task experience on the processing of extraneous information. Finally, at the end of this Chapter I will present the research questions and organization of this dissertation.

The Cognitive Theory of Multimedia Learning and Cognitive Load Theory

The CTML (Mayer, 2014) and CLT (Sweller et al., 2011) are two of the most influential theories on how humans learn from multimedia learning materials. Both argue that the human cognitive architecture, more specifically, the limitations of our working memory, should be taken into account when designing learning materials. Working memory can be defined as "a limited capacity [brain] system allowing the temporary storage and manipulation of information necessary for such complex tasks as comprehension, learning and reasoning" (Baddeley, 2000, p. 418; text in square brackets added). Working memory is limited in duration and capacity (e.g., Baddeley, 2000; Barrouillet & Camos, 2007; Cowan, 2001; Miller, 1956). With regard to the capacity limitation, Cowan (2001) proposed that our memory span is limited to around four chunks, where a chunk is a collection of items that is remembered together. Regarding the limited duration, Barrouillet and Camos (2007) proposed that working memory resources have to be shared between maintenance of 'old' information (prior knowledge from long-term memory or previously processed information during task performance), and processing of new, incoming information. When more old information elements have to be maintained active and new elements need to be processed faster, the working memory load becomes higher. Therefore, a task that is cognitively undemanding when time is unlimited can become very demanding when time is limited.

According to CTML and CLT, learning occurs when schemas are constructed or elaborated in working memory and stored in long-term memory (schema acquisition and elaboration; Sweller, 1994). Schemas are cognitive constructs that organize multiple elements of information into a single element with a specific function (Sweller, 1994; Paas, Renkl, & Sweller, 2003). Schema construction requires the selection of information from the environment (e.g., the text and picture) by attending to it and then organizing and integrating this information (with prior knowledge) into a coherent schema in working memory (Mayer, 2014). Existing schemas (i.e., prior knowledge) can be brought from long-term memory into working memory and can then be elaborated or refined with the new information (i.e., schema elaboration; Sweller, 1994). Therefore, learning new information imposes a load on working memory. Three types of cognitive load can be distinguished: 1) Intrinsic cognitive load, resulting from essential information processing, 2) extraneous cognitive load, resulting from extraneous (i.e., irrelevant/unnecessary) information processing, and 3) germane cognitive load, caused by generative information processing aimed at making sense of the instructional materials (Mayer, 2014; Paas et al., 2003).

Intrinsic cognitive load is determined by the complexity or element interactivity of the learning material (which in turn depends on the level of prior knowledge of a learner, e.g., Chandler & Sweller, 1991). Learning materials with low element interactivity (i.e., low intrinsic load) consist of elements that can be learned in isolation, without reference to other elements in the task (Sweller, 2010; Sweller et al., 2011). For example, learning vocabulary is low in element interactivity, as each new word can be remembered without reference to another word. In contrast, learning materials in which the various elements are related and must be processed simultaneously in working memory are high in element interactivity (i.e., high in intrinsic load). For instance, instructional materials about complex biological processes are high in element interactivity, as relations between information elements need to be processed and combined in schemas for learning to occur.

Extraneous cognitive load is imposed by processes that arise from the manner in which the information is presented, but do not contribute to (or may even interfere with) learning. For instance, in the context of this dissertation, extraneous load is imposed when learners are presented with irrelevant or unnecessary information and engage in processing that information. This draws on limited working memory resources without contributing to learning the essential information. Finally, germane cognitive load is imposed by generative processes that are conducive to schema acquisition and elaboration. For example, prompting learners to self-explain the principles behind biological processes that they are reading about in multimedia materials (cf. Chi, DeLeeuw, Chiu, & Lavancher 1994) will increase the demands on working memory, but improve learners' understanding.

The three types of cognitive load are additive, and the total cognitive load cannot exceed working memory resources for learning to occur. Instructional materials should therefore be designed in such a way that extraneous processing is kept to a minimum, so that all available resources can be devoted to essential and generative processing (Mayer, 2014; Sweller et al., 2011). Consequently, both the CTML and CLT state that the presentation of extraneous information, which is information that is irrelevant or unnecessary for learning, should be avoided because it hinders rather than helps learning (the coherence principle, see Mayer & Fiorella, 2014; and the redundancy principle, see Kalyuga & Sweller, 2014).

The Negative Effect of Extraneous Information on Learning

The negative effect of extraneous information on learning arises because students attend to, process, and attempt to integrate this information with the essential information. These are extraneous processes (imposing extraneous load), which unnecessarily deplete valuable working memory resources that can no longer be devoted to processing the essential information. Such extraneous processing may not be detrimental for learning when working memory capacity limits are not exceeded, for instance with materials low in intrinsic load (i.e., containing few interacting elements), or when there is sufficient time available to compensate for the extraneous processing. However, it will start to hamper learning under conditions of high intrinsic load, for example when materials are complex (with many interacting elements), or when time is constrained (Barrouillet & Camos, 2007;

Sweller et al, 2011). Consequently, it is considered important to avoid the presentation of extraneous information (cf. the coherence principle, see Mayer & Fiorella, 2014; and the redundancy principle, see Kalyuga & Sweller, 2014). Extraneous information can be either irrelevant, or unnecessary for learning.

Irrelevant Information

Information that is irrelevant for learning has no relation with the learning goal, and hampers learning when added to instructional materials. The negative effect of irrelevant information on learning has been demonstrated for instance, when instructional materials are enriched with interesting and entertaining (yet irrelevant) information (i.e., seductive details; Harp & Mayer, 1998, Mayer, Heiser, & Lonn, 2001; Moreno & Mayer, 2000; Park, Moreno, Seufert, & Brünken, 2011; Rey, 2014; Sanchez & Wiley, 2006). For example, Harp and Mayer (1998) had participants learn about the formation of lightning using a booklet with or without seductive text and pictures. These seductive details hampered recall and transfer performance compared to materials without the seductive details. A negative effect on learning has also been found when remotely related information that actively interferes with processing of the currently essential information is added to the learning materials (Mayer, DeLeeuw, & Ayres, 2007). Mayer et al. (2007) had participants learn about the working mechanisms of hydraulic brakes, with or without explanations added about caliper brakes and air brakes. Adding explanations about caliper breaks and air brakes – although related to the central content - interfered with learning the working mechanisms of hydraulic brakes, mostly when these extra explanations were presented after learning about hydraulic brakes. Finally, presenting irrelevant pictorial information (animations) that mismatches the relevant textual information has been shown to hamper learning (Hald, Van den Hurk, & Bekkering, 2015).

All in all, the negative effect of different kinds of irrelevant information (i.e., seductive, interfering, and mismatching information) on learning seems to be quite robust and does not seem to depend on whether this irrelevant information was presented as text (e.g., Rey, 2014), animation (e.g., Hald et al., 2015), or sounds (e.g., Moreno & Mayer, 2000). However, it is known that when learners have the means to compensate for the irrelevant information, this negative effect of irrelevant information on learning is smaller or absent. For instance, when learners have more attentional control (cf. Rey, 2014), they might not attend as much to the irrelevant information, and when they do not process it, it will not burden working memory. When learners have more working memory resources available (cf. Sanchez & Wiley, 2006), for instance when they have a higher working memory capacity, it is less likely that working memory is overloaded. In both cases, CLT would not predict that learning is hampered.

Unnecessary Information

Unnecessary information is related to the learning goal, but not necessary for learning (i.e., schema acquisition or elaboration) because the information is presented twice or it is unnecessarily elaborate. For instance, the negative effect of simultaneously presenting the same text in both written and spoken form has been documented quite extensively (e.g., Craig, Gholson, & Driscoll, 2002, Jamet & Le Bohec, 2006; Kalyuga, Chandler, & Sweller, 1999; Mayer et al., 2001; but see Mayer & Johnson, 2008; Yue, Bjork, & Bjork, 2013). For example, Kalyuga et al. (1999) showed that students learned better, while investing less mental effort, from a diagram with narrated text than from a diagram with narrated and on-screen text. A negative effect on learning has also been found when self-containing diagrams are accompanied by textual explanations (Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1991; Pociask & Morrison, 2008). For instance, Chandler and Sweller (1991) showed that the addition of unnecessary text to diagrams hampered learning, especially when the text was physically integrated in the diagram, or when participants were instructed to mentally integrate the text and the diagram. Finally, learning is hampered when unnecessary details and examples are added to learning materials while a more concise summary would have been sufficient (e.g., Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Reder & Anderson, 1982). For example, Reder and Anderson (1982) showed that students learned better from a textbook summary than from the original text (even when the time learners could spent on the main points was kept equal; suggesting that unnecessary details actually interfere with learning the main points of a text).

Overall, the negative effect of different types of unnecessary information (i.e., presenting identical text in written and spoken form, presenting identical information in text and diagram form, or adding unnecessary details and examples to text) on learning seems quite consistent. However, when participants were not forced to attend to the extraneous information, that is, when it was not physically integrated with essential information and when learners were not instructed to integrate it (cf. Bobis et al., 1993; Chandler & Sweller, 1991), the negative effect seems to be smaller, suggesting that participants were able to ignore the extraneous information to some extent. This in line with eye-tracking research, which suggests that with increasing task experience, people learn to focus their attention more on task-relevant information and to ignore task-irrelevant (or extraneous) information.

Learning to Ignore Extraneous Information

Several eye tracking studies have shown that with increasing task experience, people learn to focus on task-relevant and ignore task-irrelevant information. For example, Charness, Reingold, Pomplun, and Stampe (2001) showed that chess experts had different gaze patterns than intermediate chess players, making fewer fixations and fixating more on relevant pieces. Van Gog, Paas, and Van Merriënboer (2005) found that participants with more expertise in electrical circuits troubleshooting had shorter mean fixation duration and fixated more on task-relevant components of the electrical circuit in the first phase of troubleshooting, compared to participants with less expertise. Jarodzka, Scheiter, Gerjets, and Van Gog (2010), using a dynamic task, also found that experts attended more to the relevant parts of a stimulus compared to novices. However, these studies focused on expertise, that is, between-subjects differences in knowledge or skill as a function of task experience, rather than within-subjects effects, so they could not show a causal relationship between knowledge and viewing behaviour.

This causal relationship has been investigated in studies in which participants build up knowledge of the task while their viewing behavior is tracked. These within-subjects studies confirm that, even after relatively little practice, participants start to ignore task-irrelevant information and focus more on task-relevant information. For example, Haider and Frensch (1999; see also Haider & Frensch 1996) used an alphabetic string verification task and showed that participants implicitly learned to ignore the task-irrelevant information when they became more experienced with the task. They called this the *information-reduction hypothesis,* which states that, with task practice, participants first learn to distinguish task-relevant from task-irrelevant information, and then learn to focus more on the task-relevant information. Furthermore, Canham and Hegarty (2010; see also Hegarty et al., 2010) used a task in which participants had to make inferences from a weather map, either with only task-relevant information or both task-relevant and task-irrelevant information. The presence of task-irrelevant information hampered performance. However, after participants got a short instruction (10-15 minutes) about relevant meteorological principles, they spent more time viewing the relevant parts of the weather map.

These results all indicate that, during task performance, people are able to ignore task-irrelevant information as a result of their (increased) expertise on a task. The question is whether these results would be generalizable to learning materials with extraneous information. When students would also be able to start ignoring extraneous information with increasing experience, then it would no longer capture attention and working memory resources, and therefore, the negative effect of extraneous information should decrease or no longer occur with increasing task experience. This leads to the main research questions, which will be discussed in the next section.

Research Questions

The aim of this dissertation is to provide an answer to the following questions: 1) Does the negative effect of extraneous information on learning decrease or disappear with increasing task experience? 2) Does this effect arise because learners start to ignore the extraneous information? These two questions are addressed both for irrelevant and unnecessary information presentation. An important aspect of this dissertation is the use of eye-tracking methodology to address the second question, which can provide insight into the perceptual and cognitive processes that underlie the effects of different multimedia materials on learning outcomes (Van Gog & Scheiter, 2010).

Organization of this Dissertation

The two main research questions were addressed in in four empirical chapters, presented in two parts. The studies in **Part 1**, presented in Chapters 2 and 3, investigated whether the negative effect of *irrelevant* information on learning would decrease or disappear with increasing task experience. **Chapter 2** describes three experiments on the effects of irrelevant pictures when learning action-word definitions. In all experiments, participants learned the definitions of new words (from an artificial language) that denoted actions, coupled with matching pictures (depicting the same action), mismatching pictures (depicting another action), or without pictures. Experiments 1a/1b addressed the question whether adding irrelevant information (the mismatching pictures) would hamper learning of word definitions compared to learning words with matching pictures or without pictures. Experiment 2 examined the main hypothesis that irrelevant information would no longer hamper learning once learners gained experience with the word-learning task. Finally, Experiment 3 employed eye tracking to test the hypothesis that learners would adapt their study strategy and start to ignore the irrelevant information with increasing task experience.

The study presented in **Chapter 3** built on the findings from Chapter 2, by investigating whether learners suppressed attention to the content of the mismatching pictures, or 'only' learned to ignore the location of the mismatching pictures. This question was addressed in Experiments 1a/1b, by systematically changing the location of matching and mismatching pictures for half of the participants after they had accumulated task experience. If participants would only suppress attention to the picture *location*, then word learning should be negatively affected after the location change for participants engaging in word learning with mismatching pictures. However, if they were aware that the *content* of the mismatching pictures was irrelevant for learning the word definitions, they would be expected to actively suppress attention to the pictures regardless of the location, in which case performance should not be negatively affected after the location change for participants engaging in word learning with mismatching pictures.

The studies in **Part 2**, presented in Chapters 4 and 5, investigated the negative effect of *unnecessary* information on learning, using more complex (i.e., higher element interactivity) multimedia materials (i.e., expository text with pictures on biological processes). In the two experiments in **Chapter 4**, the hypothesis was addressed that learners would start to ignore unnecessary textual information (which merely described the picture) with increasing task experience, thereby reducing its negative effects on learning. In addition, it was investigated whether the layout of the unnecessary information (integrated in or separated from the picture) would influence the effect of unnecessary text on attention and learning over time (i.e., with task experience). Participants learned about the process of mitosis with materials consisting of a combination of essential text and pictures (control), essential text and pictures with unnecessary text presented either integrated in or separated from the picture. It was hypothesized that an initial negative effect of unnecessary information would occur; that this negative effect would decrease (or even disappear) as participants gained task experience; and that this decrease would be stronger when the unnecessary text was presented separated from the picture (i.e., separated unnecessary text would be easier to ignore than integrated unnecessary text).

In the two experiments described in **Chapter 5** participants studied multimedia materials about the functioning of the heart. This study involved a replication of

Experiment 5 of Chandler and Sweller (1991), which was part of a series of experiments showing that (1) the addition of unnecessary text to a self-containing diagram impeded learning; (2) this negative effect of unnecessary text was larger when the text was spatially integrated in the essential materials as compared to spatially separated; and (3) this effect was larger when participants were instructed to mentally integrate the separated unnecessary information as compared to no such instruction. A second aim of the first experiment in Chapter 5 was to examine the influence of pacing (system-paced vs. self-paced) on the occurrence and size of the negative effect of unnecessary information, as this negative effect might be larger when learning is system-paced, because under system-paced conditions, any unnecessary information processing goes directly at the expense of time available for essential information processing.

In these experiments, we used four different lay-outs in four different conditions: 1) a diagram presented without unnecessary text (diagram only); 2) a diagram with unnecessary text separated from the diagram (separated); 3) a diagram with unnecessary text separated from the diagram with the instruction to mentally integrate the text and the diagram (integration instruction); or 4) a diagram with unnecessary text integrated into the diagram (integrated). In Experiment 1 study time was self-paced for half of the participants, and system-paced for the other half. We hypothesized to find a negative effect of the unnecessary text when it was integrated in the essential material, or when participants received the integration instruction. We also addressed the open question whether participants would be able to ignore the spatially separated unnecessary text to such an extent that it would not hamper their learning compared to a diagram-only condition. In addition, we hypothesized that system-pacing would aggravate the negative effects of unnecessary text on learning, particularly in the integrated condition and the integration-instruction condition because these conditions were assumed to impose the highest cognitive load on the learner.

Finally, **Chapter 6**, presents a summary and general discussion of the main results of this dissertation.

Chapter 2

Task Experience as a Boundary Condition for the Negative Effects of Irrelevant Information on Learning

This chapter has been published as:

Rop, G., Van Wermeskerken, M., De Nooijer, J. A., Verkoeijen, P. P. J. L., & Van Gog, T. (in press). Task experience as a boundary condition to the negative effects of irrelevant information on learning. *Educational Psychology Review*. doi: 10.1007/s10648-016-9388-

Abstract

Research on multimedia learning has shown that learning is hampered when a multimedia message includes extraneous information that is not relevant for the task, because processing the extraneous information uses up scarce attention and working memory resources. However, eye-tracking research suggests that task experience might be a boundary condition for this negative effect of extraneous information on learning, because people seem to learn to ignore task-irrelevant information over time. We therefore hypothesized that extraneous information might no longer hamper learning when it is present over a series of tasks, giving learners the chance to adapt their study strategy. This hypothesis was tested in three experiments. In Experiments 1a/1b, participants learned the definitions of new words (from an artificial language) that denoted actions, with matching pictures (same action), mismatching pictures (another action), or without pictures. Mismatching pictures hampered learning compared to matching pictures. Experiment 2 showed that task experience may indeed be a boundary condition to this negative effect on learning: The initial negative effect was no longer present when learners gained experience with the task. This suggests that learners adapted their study strategy, ignoring the mismatching pictures. That hypothesis was tested in Experiment 3, using eye tracking. Results showed that attention to the pictures waned with task experience, and that this decrease was stronger for mismatching than for matching pictures. Our findings demonstrate the importance of investigating multimedia effects over time and in relation to study strategies.

Introduction

Multimedia learning, which can be defined as learning with a combination of words (written or spoken) and pictures (static or dynamic), has been widely investigated in research inspired by the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2014) and Cognitive Load Theory (CLT; Sweller, Ayres, & Kalyuga, 2011). This has led to the establishment of several principles for designing effective multimedia instructions. The present study is concerned with the coherence principle, which states that presenting extraneous information that is not relevant for the learning task should be avoided, because it hinders rather than helps learning (Mayer & Fiorella, 2014). Because eye-tracking research has shown that with increasing task experience, people learn to ignore irrelevant information during task performance, we hypothesized that task experience might be a boundary condition for the negative effect of extraneous information on learning. That is, the negative effect that the presentation of extraneous information initially has on learning might no longer occur when this information is present (in the same location) over a series of tasks, because learners might adapt their study strategy (i.e., learn to ignore the extraneous information). This hypothesis was tested in a series of three experiments, which will be introduced after discussing the relevant literature in more detail.

Cognitive Load in Multimedia Learning

Decades of multimedia research have shown that learning often improves when study tasks or materials combine pictorial and verbal representations of the content (i.e., the multimedia effect; Butcher, 2014). However, it soon became apparent that there are circumstances under which this hampers rather than aids learning (e.g., Chandler & Sweller, 1991; Harp & Mayer, 1998). These circumstances are related to the limitations of working memory, which can be defined as “a limited capacity [brain] system allowing the temporary storage and manipulation of information necessary for such complex tasks as comprehension, learning and reasoning” (Baddeley, 2000, p. 418; text in square brackets added). Working memory is limited in both duration and capacity (e.g., Baddeley, 2000; Barrouillet & Camos, 2007; Cowan, 1995; Miller, 1956). For instance, on average, our memory span is ‘seven plus or minus two’ chunks, where a chunk is one piece of information (Miller, 1956). Barrouillet and Camos (2007) propose that the limited working memory resources have to be shared by rapidly switching attention between maintenance of ‘old’ information (prior knowledge from long-term memory or previously processed information during task performance), and processing of new, incoming information. Consequently, the higher the number of old information elements that have to be maintained active, and the faster new elements need to be processed, the higher the working memory load.

Learning (i.e., schema construction/elaboration in long-term memory; Sweller, 1994) requires that old information is maintained active in working memory and successfully integrated with the new information presented in the learning materials. When these processes are disrupted, learning is hampered. Moreover, learning may be hindered when scarce working memory resources are devoted to processing extraneous information that is not necessary for the learning task. Such extraneous processing may not be detrimental for learning when capacity limits are not exceeded, for instance with simple materials (that contain few interacting information elements), or when there is sufficient time available to compensate for the extraneous processing. However, it will start to hamper learning when materials are complex (with many interacting information elements), or when time is constrained (Barrouillet & Camos, 2007; Sweller et al, 2011). Consequently, it is important to avoid the presentation of extraneous information that does not contribute to learning as much as possible.

Avoiding the Presentation of Extraneous Information in Multimedia Learning

That the presentation of extraneous information can have a negative effect on learning has been established in many experiments. However, there are two different types of extraneous information presentation effects. The first, which is generally called the redundancy effect (Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014; Sweller et al., 2011), concerns the negative effect of the presentation of *identical* extraneous information to learners in two modalities compared to a single modality. For example, it has been shown that when the text accompanying pictures or animations is presented simultaneously in both spoken and written form, this hampers learning compared to spoken text only (e.g., Craig, Gholson, & Driscoll, 2002; Mayer, Heiser, & Lonn, 2001; but see Mayer & Johnson, 2008; Yue, Bjork, & Bjork, 2013).

The second, which is called the coherence effect in CTML (Mayer & Fiorella, 2014; Mayer & Moreno, 2003), concerns the negative effect on learning of the presentation of extraneous information that is not relevant or necessary for learning, but is added to enrich or elaborate learning materials, compared to when this is left out. For instance, learning is hampered when interesting and entertaining information that is related to the topic but irrelevant for the learning task at hand is added to enrich materials (i.e., seductive details, e.g., ‘fun’ facts, pictures, videos, or sounds; Harp & Mayer, 1998; Mayer et al., 2001; Moreno & Mayer, 2000); when learning materials are unnecessarily elaborate, presenting textual explanations with self-explanatory diagrams (Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1991) or presenting details and examples whereas a concise and coherent summary would suffice (e.g., Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Reder & Anderson, 1982); or when information on related systems is presented when learning about a specific system (Mayer, DeLeeuw, & Ayres, 2007).

The negative effects of extraneous information presentation on learning presumably arise because learners attend to, process, and attempt to integrate the extraneous information with the essential information, which unnecessarily depletes valuable working memory resources. Moreover, in some cases of the coherence effect, the content of the additional information that is presented may actively interfere with learning the essential information. For instance, Mayer et al. (2007) showed that adding explanations about caliper breaks and air brakes interfered with learning the working mechanisms of hydraulic brakes. Participants who learned about the caliper and air brakes made more intrusion errors (i.e., including information about caliper and air breaks in their answers) than participants who only learned about hydraulic breaks. The present study addresses the coherence effect, by investigating the effects of extraneous pictorial information on word learning in a 'second' (artificial) language.

Multimedia and Coherence Effects in Word Learning

As mentioned above, the multimedia effect refers to the finding that learning often improves when study tasks or materials combine pictorial and verbal representations of the content (Butcher, 2014). In this case, additional information is also presented, but it is not extraneous to (i.e. does not hamper) learning and may even facilitate learning.

With regard to word learning, some studies have shown that adding pictures of the word to be learned (and example sentences in which the word is used), does not hamper and -in accordance with the multimedia effect- can even foster word learning (for a review, see Sadoski, 2005). For example, Smith, Stahl, and Neil (1987) taught undergraduate students novel words in their first language (English) in one of three conditions: Definition only; definition and a sentence using the word; and definition, a sentence, and a picture. Results on the immediate retention test favoured the condition with pictures, although the differences were not significant. On a delayed retention test, the condition with pictures performed significantly better than the definitions only condition (but not than the definitions plus sentences condition), demonstrating a multimedia effect (i.e., pictures facilitating word learning; Butcher, 2014).

Whether a multimedia (i.e. facilitative) effect of pictures on word learning is found, however, may depend on how easily the words can be mentally simulated or visualized. For instance, in a second language vocabulary learning study, Farley, Ramonda, and Liu (2014) taught Spanish vocabulary to English-speaking university students and found that abstract words (i.e., words without a physical referent) were learned better with pictures, while there was no such effect for concrete words (i.e., words with a physical referent). Concrete words may be easy to mentally visualize and learn because of the physical referent (e.g., Altarriba & Bauer, 2004), in which case pictures do not have much added value for learning. Shen (2010) found a comparable effect in learning Chinese as a foreign language: Pictures improved word learning, but only for abstract words. For non-abstract words, even though presenting pictures did not help, it did not hinder learning either.

While additional presentation of pictures of the words to be learned, might not hamper and could even foster learning, presentation of pictures that do not match the words might interfere with learning. For instance, a recent study on effects of animations on action word learning (e.g., to chisel, to hoe) in the first language (Hald, Van den Hurk, & Bekkering, 2015), included a mismatched animation condition, to control for effects of movements shown in the animations. Results showed that word learning was significantly hampered in the animation condition in which the actions depicted in the animation mismatched the word to be learned compared to when it matched the word to be learned.

Another study investigated the effects of pictures on action word learning in a "second" artificial language, which matched or mismatched the learners' handedness (De Nooijer, Van Gog, Paas, & Zwaan, 2013). Participants first saw the artificial language word (e.g. 'luko') on the screen, and heard a verbal definition of the word (e.g. 'luko' means 'to dispense from a container'). Hearing actions described tends to automatically result in a body-specific (Casasanto, 2009) mental simulation/visualization of the action (e.g., Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermüller, Hauk, Nikulin, & llmoniemi, 2005; Tettamanti, et al., 2005). Then, they heard this definition a second time, but now a picture was shown along with the artificial language word. This picture always showed the defined action, but either matched (i.e., right-handed for right-handers) or mismatched (i.e., left-handed for right-handers) participants' mental simulation. The mismatching pictures hampered right-handers' learning (this was a small but consistent effect across multiple experiments)[1], presumably because attending to and processing the picture that mismatched with their mental simulation of the action, interfered with learning the verbal definition.

In sum, additional presentation of pictures that match the word to be learned might not hamper, and may even help word learning. In contrast, pictures (whether static or dynamic) that mismatch with (the mental visualization of) the word to be learned have been shown to hinder learning. It is assumed that mismatching pictures hamper word learning because they capture learners' attention, and lead them to engage in extraneous –and conflicting- information processing. However, it is unclear whether extraneous information would continue to hamper learning when it is present (in the same location) over a series of tasks. This would give learners the chance to adapt their study strategy and ignore the extraneous information. Indeed, eye-tracking research suggests that with increasing task experience, people learn to focus their attention more on task-relevant information and to ignore task-irrelevant information.

Learning to Ignore Task-Irrelevant Information

Several studies have shown that with increasing task experience, people learn to focus on task-relevant and ignore task-irrelevant information. For instance, expertise research showed that chess experts had different viewing patterns compared to intermediate chess players, which included making fewer fixations and fixating more on relevant pieces (Charness, Reingold, Pomplun, & Stampe, 2001). Likewise, Van Gog, Paas, and Van Merriënboer (2005) demonstrated that participants with more expertise in an electrical circuit-troubleshooting task had shorter mean fixation duration and fixated more on task-relevant components of the electrical circuit in the first phase of troubleshooting, compared to participants with less expertise. Furthermore, Jarodzka, Scheiter, Gerjets, and Van Gog (2010) showed that, in a visually complex dynamic task, experts attended more to the relevant parts of a stimulus compared to novices. While these studies suggest that people are able to ignore irrelevant information with increasing expertise, they did not show a causal

[1] Left-handers' learning, in contrast, was not affected by right-handed pictures, presumably because they have visual and actual experience with the right hand being used for these actions.

relationship between task experience and viewing behaviour, as they compared existing groups of experts and novices or intermediates.

Such evidence also exists, however, and comes from studies that investigated how viewing behavior changed as expertise developed. These studies confirm that, even after relatively little practice, participants start to ignore task-irrelevant information and focus more on task-relevant information. For example, Haider and Frensch (1999) used an alphabetic string verification task and showed that participants implicitly learned to ignore the task-irrelevant information when they became more experienced with the task. Furthermore, Canham and Hegarty (2010; see also Hegarty, Canham, & Fabrikant, 2010) used a task in which participants had to learn to make inferences from weather maps, and showed that participants fixated more on task relevant information after a short training (10-15 minutes) than before a training about relevant meteorological principles.

Taken together, these findings strongly suggest that people may learn to ignore irrelevant information during task *performance*, as a result of increasing experience with a task. However, these studies did not investigate effects on *learning*. When learners are able to start ignoring extraneous information with increasing experience during learning, then it would no longer capture attention and working memory resources, and therefore, the negative effect of extraneous information should decrease or no longer occur with increasing task experience. In other words, these findings suggest that task experience might be a boundary condition to the negative effect of extraneous information on learning. This is not only theoretically relevant to establish, as it provides more insight into the underlying cognitive mechanisms of multimedia effects and the role that study strategies play over time, but it is also relevant for instructional designers, because some multimedia principles are hard to implement generally (e.g. what may be essential information for someone with low prior knowledge, may be extraneous for a more advanced learner). The present study addressed this hypothesis in a series of three experiments, using a word-learning task.

The Present Study

Experiments 1a and 1b were conducted to first establish whether mismatching pictures would negatively affect word learning compared to no pictures and matching pictures. We used a similar experimental design as De Nooijer et al. (2013), but with right-handed pictures and right-handed participants only. Participants learned words from an artificial language called Vimmi (Macedonia & Knösche, 2011) that we coupled with action verb definitions. We used an artificial language to exclude any influences of prior knowledge of the words to be learned and associations on idiomatic level between two languages. Participants saw the artificial language word on the screen, and heard the verbal definition of the word (e.g., 'ifra' means 'to polish or scrape with sandpaper'), after which they heard this definition a second time, either without a picture (control condition), with a matching picture (showing the action) or a mismatching picture (showing another action). Note that matching pictures, even though they showed the action that was being defined, are not entirely redundant to the definition and may in fact provide useful additional information. For instance, in case of known actions, these pictures may automatically prompt participants to think of the English verb that was not part of the definition (e.g., in the case of 'ifra', the picture might clarify that they are hearing the definition of 'to sand'), and in case of unknown actions, the pictures might help understand the meaning of the word.

We expected that the mismatching pictures would capture attention and interfere with processing the verbal definition the second time it was presented, which would hamper learning compared to the no pictures condition (i.e., a coherence effect) and the matching pictures condition (as we expected this condition to do as well or better than the no picture condition, see below). Learning could be hampered by mismatching pictures via two (not mutually exclusive) routes: 1) less attention could be devoted to processing the materials relevant for learning, and 2) processing the content of the mismatching action pictures could actively interfere with learning the correct action word definition. With respect to the matching pictures, we expected based on the studies described in the introduction that these would either improve learning the meaning of new words (Farley et al, 2014; Shen, 2010; Smith et al, 1987), providing evidence of a multimedia effect, or not affect word learning. If they do not significantly contribute to word learning, one could argue that matching pictures also constitute extraneous information, in the sense that they are redundant and could be left out at no expense to learning. As such one could argue that this null finding would also be a kind of coherence effect, though not in the strict definition (Mayer & Fiorella, 2014), as excluding the pictures would not have a positive effect on learning compared to including them.

To foreshadow, Experiment 2 addressed the main question of whether the negative effect of mismatching pictures on word learning would decrease or no longer occur with increasing task experience, which would suggest that learners adapted their study strategy. Experiment 3 subsequently investigated study strategies directly, by means of eye tracking (to measure attention allocation), to determine whether participants indeed learned to ignore the mismatching pictures over time, with task experience.

Experiments 1a and 1b

As described above, Experiments 1a and 1b were conducted to establish whether mismatching pictures would negatively affect word learning compared to no pictures and matching pictures. Experiment 1b was a direct replication of Experiment 1a to test the reliability of our results.

Method

Participants and design

Participants (Experiment 1a: n = 85, Experiment 1b: n = 144) were recruited via Amazon's Mechanical Turk (Paolacci, Chandler, & Ipeirotis, 2010), and were paid 0.75 dollar for their participation[2]. A-priori defined criteria for post-hoc exclusion were the following: Being a non-native English speaker (n = 2 and n = 1, resp.); being left-handed (i.e., the pictures were right-handed and even though left-handers seem less hampered by right-handed pictures than right-handers are by left-handed pictures according to findings by De Nooijer et al., 2013, we wanted to rule out any potential handedness effects; n = 9 and n = 21, resp.); participating in the experiment twice or having participated in a similar earlier experiment (n = 2 and n = 3, resp.); and being in a noisy environment (i.e., self-reported noise of seven or higher on a scale of one to nine; n = 4 and n = 7, resp.). Thus, for Experiment 1a our final sample comprised 68 participants (M_{age} = 39.13 years, SD = 14.12 years, range 20-74; 47 females), and for Experiment 1b our final sample comprised 112 participants (M_{age} = 34.58 years, SD = 12.66 years, range 19-89; 78 females). Both experiments employed a within-subjects design, so participants learned words under all three conditions.

Materials

Participants learned 18 Vimmi words presented in Qualtrics software (Qualtrics, Provo, UT). Each word was randomly coupled to the definition of an action verb (e.g., 'ifra' means 'to polish or scrape with sandpaper'). In every condition, the participants saw the Vimmi word and heard the definition (spoken by a female voice; M_{length} = 3.44 seconds, SD = 1.01 seconds) of the word they had to learn twice, each presentation lasted 11 seconds and the program automatically progressed. In the two

[2] Note that this was a common level of payment at the time the experiments were conducted; we are aware of the recent discussions of and increases in MTurk wages.

picture conditions, a matching picture (showing the action) or a mismatching picture (showing another action) accompanied the word (see Figure 1) the second time participants heard the definition. Participants' knowledge of the definition was tested with a cued recall test, in which they were presented with the written Vimmi word and had to type in the definition as literally as possible.

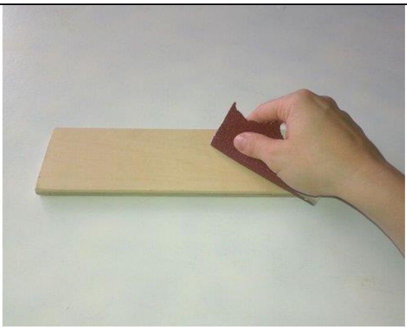
Word	Definition	Condition	Picture
Ifra (to sand)	To polish or scrape with sandpaper	Matched	
		Mismatched	

Figure 1. Example materials. The spoken definition is presented twice, the second time accompanied by a picture in the matched condition (top row) and mismatched condition (bottom row).

Procedure

Participants learned the words in three blocks of six, and after each block the cued recall test for those words was administered. During the cued recall test no time constraint was imposed. To avoid confusion, the six words in the control (no picture) condition were presented in one block (either the first or the last) because participants might think a technical error occurred when no picture would be present at random words during the experiment. The other two blocks consisted of three words with matching and three words with mismatching pictures. Order within block was randomized, so participants did not know beforehand if the next picture would be matching or mismatching the definition, and, thus, could not anticipate on the usefulness of the picture. Furthermore, the artificial words were rotated across definitions to control for possible item effects, resulting in the use of eight lists. In total, the experiment lasted about twenty minutes and was administered without breaks.

Scoring

In all experiments, participants were awarded 1 point if the complete definition was given on the cued recall test. When a part of the definition was missing, they received 0.5 point. If they did not provide a definition, or if it was completely wrong, 0 points were awarded. So, every participant could score a maximum of six points on each test. We adopted the scoring scheme of De Nooijer et al. (2013) who used the same materials (they found an interrater reliability of $\kappa = .82$), but established interrater reliability again for the present study, by having part of the data (10.6%) scored by a second, independent rater. Both raters saw only the definitions and words during scoring, so they were blind to the experimental condition under which each word was learned. Because interrater reliability was sufficient ($\kappa = .76$; 'substantial agreement' according to Landis & Koch, 1977), the scores from the first rater (first author) were used in this and all subsequent experiments.

Results

Throughout all experiments, we maintained an alpha level of .05, and when the sphericity assumption was violated, we reported the Greenhouse-Geisser correction. Effect size measures used were partial eta-squared and Cohen's *d*. Both can be interpreted in terms of small ($\eta_p^2 \sim .01$, $d \sim 0.2$), medium ($\eta_p^2 \sim .06$, $d \sim 0.5$), and large ($\eta_p^2 \sim .14$, $d \sim 0.8$) effect sizes (Cohen, 1988). In addition, effect sizes can also be interpreted with respect to median effects sizes found for these effects (in other studies with different materials). For example, Mayer and Fiorella (2014) report an average effect size of $d = 0.86$ for the coherence effect (based on $n = 23$ comparisons). In Experiments 1a and 1b, data were analysed with a repeated-measures ANOVA with Condition (matched, mismatched, control) as a within-subjects variable and the scores on the cued recall test as dependent variable.

Table 1 shows the means and standard deviations for the cued recall tests of Experiments 1a and 1b for the different conditions. In Experiment 1a there was a main effect of Condition, $F(1.83, 122.26) = 3.37$, $p = .042$, $\eta_p^2 = .05$. Bonferroni corrected post-hoc tests showed that the mismatched condition performed worse than the matched condition, $p = .028$, $d = 0.32$. There were no significant differences between the control and matched condition, $p > .999$, $d < 0.01$, nor between the control and mismatched condition, $p = .125$, $d = 0.25$.

Experiment 1b replicated the results of Experiment 1a, again showing a main effect of Condition, $F(1.82, 202.19) = 4.15$, $p = .017$, $\eta_p^2 = .04$, with Bonferroni corrected post-hoc tests indicating that the mismatched condition performed worse than the matched condition on the recall test, $p = .002$, $d = 0.32$, and that there were not significant differences between the control and matched condition, $p = .769$, $d = 0.11$, nor between the control and mismatched condition, $p = .371$, $d = 0.15$.

Table 1. *Mean (and SD) Recall Performance (max. = 6) as a Function of Condition in Experiments 1a and 1b.*

	Experiment 1a	Experiment 1b
Matched	3.65 (1.83)	3.56 (1.75)
Mismatched	3.24 (1.83)	3.15 (1.81)
Control	3.65 (1.78)	3.38 (1.80)

Discussion

The results of Experiments 1a and 1b provided evidence that adding mismatching pictures to word learning decreased recall performance compared to matching pictures, but there was no significant negative effect compared to no pictures (i.e. no coherence effect). Processing the matching pictures did not significantly improve word learning compared to the no picture control condition (i.e., there was no multimedia effect), but did not hurt learning either (which could be regarded as a kind of coherence effect, though not in the strictest definition, as exclusion did not lead to better learning than inclusion of the matching pictures). This may be due to the rather concrete object-manipulation verbs we used in our experiment, as previous research has suggested that pictures benefit learning abstract words more than learning concrete words (Altarriba & Bauer, 2004; Farley et al., 2014; Shen, 2010).

The fact that we did not find a coherence effect, indicating that including mismatching pictures would lead to poorer learning outcomes than excluding mismatching pictures (control condition) limits our conclusions somewhat. However, we did find a negative effect of mismatching compared to matching pictures on word learning. This effect of mismatching vs. matching pictures was small, and although it is smaller than the median effect size from other studies on the coherence effect, using different materials (*d* = 0.86 as reported by Mayer & Fiorella, 2014), it was comparable to the effects found by De Nooijer et al. (2013) who used the same materials but with mismatches in terms of handedness. Furthermore, the effect sizes were consistent across Experiments 1a and 1b. As this difference between mismatching and matching picture conditions is still interesting for our purpose of investigating whether task experience affects processing of mismatching pictures and thereby, learning outcomes, we conducted Experiment 2 to address this question concerning the effects of task experience. We did not find significant differences compared to the control condition for either picture condition in Experiment 1a/1b. Performance was somewhat lower in Experiment 1b than in Experiment 1a, but the pattern of results was the same. We opted to again include a no-picture control condition in Experiment 2, because we could not exclude the possibility that differences between the picture conditions and the control condition might arise over time.

Experiment 2

In Experiment 2, we used the same materials as in Experiments 1a and 1b, but now in a between-subjects design. Participants learned the words in three blocks of five, with matching, mismatching, or no pictures. We hypothesized that the negative effect of mismatching compared to matching pictures would occur initially (after the first block), but would decrease or no longer occur with increasing task experience (blocks two and three). That is, when participants would learn that the mismatching pictures are unnecessary for learning and adapt their study strategy accordingly, ignoring these pictures in the later blocks, their recall performance would increase in block 2 and 3. For the matched and control conditions we had no reason to expect that recall performance would change with increasing task experience.

We also explored how participants experienced the task[3]. To get an indication of potential differences in processing demands experienced between conditions and over time, participants were asked to rate how much mental effort they invested in

---

[3] We asked participants to rate invested effort, enjoyment, and whether they preferred to learn words without, with matching, or with decorative pictures in the future (cf. Yue et al., 2013) in Experiments 2 and 3 (this was not possible in Experiments 1a and 1b, where the blocks comprised words presented under different conditions). Only results on effort are reported; data on enjoyment (only null-effects) and preferences can be obtained from the first author.

learning the words (which is an indicator of how much cognitive load participants experienced: Paas, Tuovinen, Tabbers, & Van Gerven, 2003).

Method

Participants and design

All participants (n = 232) were recruited via Amazon's Mechanical Turk and were paid 0.75 dollar for their participation[2]. They were randomly assigned to one of the three conditions. The same exclusion criteria as in the former experiments were used: being a non-native English speaker (n = 4); being left-handed (n = 36); participating in the experiment twice or having participated in a similar earlier experiment (n = 14); and being in a noisy environment (n = 6). Finally, there were participants in the mismatched condition who did not follow the instructions[4], who were also excluded (n = 5). Thus, our final sample comprised 166 participants (M_{age} = 36.50 years, SD = 12.53 years, range 18-69; 105 females), distributed across the control (n = 62), matched (n = 55), and mismatched (n = 49) conditions.

Materials and procedure

The materials and procedure were similar to Experiments 1a and 1b, except that: (1) participants learned 15 word definitions in three blocks of five words under the same condition, (2) after each block of five words and prior to the cued recall test of that block, participants were instructed to indicate how much effort they invested in learning the words, on a nine-point rating scale ranging from 1, *very, very low effort*, to 9, *very, very high effort* (Paas, 1992), and (3) at the end of the experiment, enjoyment and instruction preference was measured (not reported, see [3]). The order of the blocks was alternated between participants using a Latin-square design, resulting in three lists per condition. The experiment lasted around twenty minutes and was administered without breaks.

Results

Unless otherwise specified, all data were analysed with a mixed ANOVA with Condition (matched, mismatched, control) as a between-subjects factor and Word block (first, second, third) as a within-subjects factor.

Test performance

Table 2 shows the scores on the cued recall test. The analysis showed no main effect of Condition, $F(2, 163) = 1.44$, $p = .239$, $\eta_p^2 = .02$, but there was a main effect of

---

[4] These participants wrote down one word that described the mismatching picture (i.e., wrench when the artificial word 'Ifra' was tested), for each and every definition they had to provide.

Word block, $F(1.85, 300.90) = 10.34$, $p < .001$, $\eta_p^2 = .06$, indicating that the scores on the cued recall test improved over the course of the experiment. Importantly, this main effect was qualified by an interaction effect, $F(3.69, 300.90) = 2.46$, $p = .050$, $\eta_p^2 = .03$.

To follow up on the interaction effect, we first analysed differences in test performance between conditions per block, using one-tailed independent-samples t-tests with a Bonferroni corrected *p*-value (i.e., multiplying the *p*-value by three, the number of tests that were performed). In line with our hypothesis, performance was significantly lower in the mismatched condition than in the matched condition in block 1, $t(102) = 2.18$, $p = .047$, $d = 0.43$, but only numerically, not significantly lower than in the control condition, $t(109) = 0.91$, $p = .545$, $d = 0.17$ (cf. Experiment 1). Finally, there was no difference in performance between the matched and the control condition in block 1, $t(115) = 1.42$, $p = .239$, $d = 0.26$ (cf. Experiment 1). In block 2 and 3, there were no significant differences between the conditions. Mismatched vs. matched: Block 2, $t(102) = 0.12$, $p > .999$, $d = 0.02$; block 3, $t(102) = 0.01$, $p > .999$, $d = 0.02$. Mismatched vs. control: Block 2, $t(109) = 1.73$, $p = .128$, $d = 0.34$; block 3, $t(109) = 1.37$, $p = .260$, $d = 0.26$. Matched vs. control: Block 2, $t(115) = 1.62$, $p = .161$, $d = 0.30$; block 3, $t(115) = 1.32$, $p = .287$, $d = 0.24$.

Table 2. *Mean (and SD) Recall Performance (max. = 5) and Mental Effort Rating (max. = 9) as a Function of Condition and Word Block in Experiment 2.*

	Performance			Mental effort		
	Block 1	Block 2	Block 3	Block 1	Block 2	Block 3
Matched	3.15 (1.46)	3.21 (1.51)	3.14 (1.63)	7.20 (1.63)	7.33 (1.63)	7.33 (1.75)
Mismatched	2.51 (1.51)	3.17 (1.45)	3.13 (1.44)	7.73 (1.15)	7.94 (1.25)	7.86 (1.47)
Control	2.77 (1.43)	3.65 (1.39)	3.51 (1.42)	7.52 (1.29)	7.85 (1.13)	7.69 (1.43)
Overall	2.82 (1.48)	3.36 (1.46)	3.27 (1.50)	7.48 (1.38)	7.70 (1.37)	7.62 (1.56)

Secondly, we probed the interaction effect by analysing the change in test performance over blocks within each condition. Repeated measures ANOVAs per condition, with Word block (first, second, third) as a within-subjects factor, showed no effect of Word block in the matched condition, $F(2, 108) = 0.08$, $p = .926$, $\eta_p^2 < .01$, a main effect of Word block in the mismatched condition, $F(2, 96) = 4.75$, $p = .011$, $\eta_p^2 = .09$, and a main effect of Word block in the control condition, $F(2, 122) = 9.86$, $p <$

.001, $\eta_p^2 = .14$. Repeated contrasts follow-up showed that, in line with our expectations, participants' performance in the mismatched condition improved from block 1 to 2, $F(1, 48) = 7.63$, $p = .008$, $\eta_p^2 = .14$, and remained stable from block 2 to 3, $F(1, 48) = 0.04$, $p = .848$, $\eta_p^2 < .01$. Similarly, though rather surprisingly, participants' performance in the control condition also improved from block 1 to 2, $F(1, 61) = 18.34$, $p < .001$, $\eta_p^2 = .23$, but not from block 2 to 3, $F(1, 61) = 0.53$, $p = .469$, $\eta_p^2 = .01$.

Mental effort

Table 2 shows the descriptive statistics of the mental effort ratings. Despite a trend in the means suggesting that participants in the matched condition invested less mental effort in the task, the analysis revealed no significant main effect of Condition, $F(2, 163) = 2.68$, $p = .072$, $\eta_p^2 = .03$[5]. There was a main effect of Word block, $F(1.90, 309.30) = 3.59$, $p = .029$, $\eta_p^2 = .02$, but no interaction effect, $F(3.80, 309.30) = 0.32$, $p = .859$, $\eta_p^2 < 0.01$. To follow-up on the main effect of Word block, repeated contrasts were conducted, which showed that invested mental effort increased from block 1 to 2, $F(1, 163) = 7.60$, $p = .006$, $\eta_p^2 = .05$, and remained stable from block 2 to 3, $F(1, 163) = 1.08$, $p = .299$, $\eta_p^2 = .01$.

Discussion

Consistent with our results from Experiment 1a/1b, the cued recall performance in the first block was lower in the mismatched condition than in the matched condition. Moreover, and in line with our hypothesis, this negative effect of mismatching pictures compared to matching pictures disappeared when participants gained more task experience. The finding that performance in the mismatched condition improved from block 1 to block 2 (and remained constant from block 2 to block 3), suggests that participants adapted their study strategy and learned to ignore the mismatching pictures.

Although this should be interpreted with caution, as the effect of condition did not reach statistical significance, an explorative analysis (see [5]) suggested that mental effort invested in the mismatched condition was higher than in the matched condition. Furthermore, participants in all conditions started to invest somewhat more effort in block 2 compared to block 1, which then remained at the same level in

[5] Although the effect of condition failed to reach statistical significance, we conducted an explorative analysis based on a remark by one of the reviewers, which showed that average mental effort invested in the mismatched condition ($M = 7.84$, $SD = 1.10$) was higher than in the matched condition ($M = 7.28$, $SD = 1.59$) when compared directly in a t-test, $t(96.18) = 2.11$, $p = .038$ (two-sided), $d = 0.41$.

block 3. Perhaps, the initial cued recall task in the first word block might have been more difficult than expected, leading participants to adjust their effort investment in the following blocks (cf. Brehm & Self, 1989; Kahneman, 1973). However, this increased effort investment only resulted into better test performance in the mismatched and control conditions, not in the matched condition (perhaps because the first block baseline score was already relatively high in this condition).

Experiment 3

The most important finding from Experiment 2 was that task experience indeed seems to be a boundary condition for the negative effect of mismatching compared to matching pictures on learning. We assume that this finding arose because participants adapted their study strategy and started to ignore the mismatching pictures in blocks 2 and 3. Experiment 3 was set up as a direct test of this assumption. Using eye-tracking methodology, we measured participants' attention allocation to the pictures in the matched and mismatched conditions over time. We hypothesized that with increasing task experience participants in the mismatched condition would allocate less attention to the pictures. We had no specific hypothesis about attention distribution in the matched condition. Because of the smaller sample size in Experiment 3 (which was sufficiently large to address our hypotheses regarding attention allocation) and the small effect size in Experiments 1 and 2, we expected to replicate the trends in performance scores from Experiment 2 (i.e., mismatched < matched in block 1; increase from block 1 to 2 in the mismatched condition), but we did not necessarily expect these to be significant.

Method

Participants and design

Participants were 96 Dutch undergraduate university students (M_{age} = 20.35 years, SD = 2.12 years; 80 female) who participated for course credit. All participants were native Dutch speakers with normal or corrected-to-normal vision. Participants were randomly assigned to either the matched or mismatched condition. Within each condition, the picture location was counterbalanced, for half of the participants the picture was presented above the word, and for the other half the picture was presented underneath the word, to rule out the possibility that the hypothesized effects would be due to a particular location.

Thirteen participants (nine in the matched and four in the mismatched condition) turned out to be left-handed and were therefore excluded from the analyses. The data of the remaining 83 participants (39 in the matched and 44 in the mismatched condition) was scored and analysed.

Apparatus and materials

The words and pictures were the same as those used in Experiment 2, but the definitions were now presented in Dutch. The Vimmi words and pictures were presented in SMI Experiment Center (Version 3.3; SensoMotoric Instruments) on a monitor with a resolution of 1680 x 1050 pixels (see Figure 2). Pictures were 500 x 400 pixels; the center of the pictures was either at 656 or at 393 pixels (above or underneath the word, respectively) on the vertical axis and at 825 pixels on the horizontal axis. Participants' eye movements were recorded using an SMI RED 250 eye tracker (SensoMotoric Instruments) that recorded binocularly at 250 Hz using iView software (Version 2.8; SensoMotoric Instruments) and subsequently analysed using BeGaze software (Version 3.3; SensoMotoric Instruments).

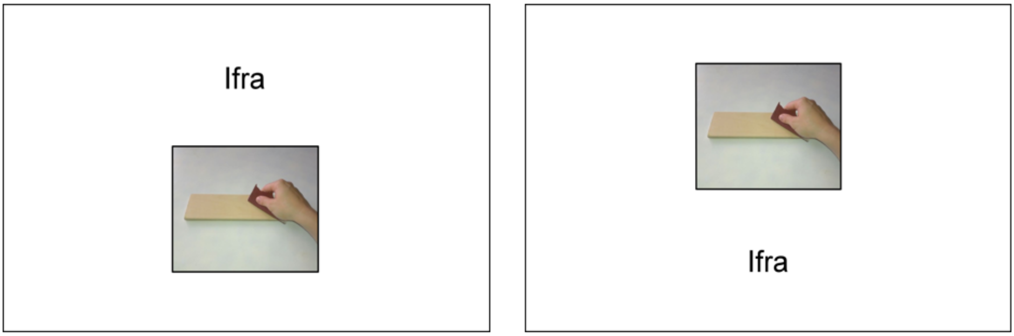


Figure 2. Example materials. Vimmi word "*Ifra*" (to sand) accompanied by the matching picture underneath (left) or above (right) the word.

Procedure

The procedure was similar to that of Experiment 2; participants learned the words in three blocks of five words and had to fill out the mental effort scale and complete a cued recall test after each block. At the start of the experiment, participants were seated in front of the monitor with their head positioned in a chin and forehead rest. The distance to the monitor was approximately 70 cm. After a short introduction, the eye tracker was calibrated using a five-point calibration plus four-point validation procedure, and participants were instructed to move as little as possible. After studying the first block of words, participants were allowed to move freely as they had to complete the recall test. After completion of the test, participants were calibrated again, after which the second block started. This procedure was repeated for block three. The order of the blocks was alternated between participants using a Latin-square design, resulting in three lists per condition. The experiment lasted around twenty minutes and was administered without breaks.

Scoring and data analysis

Due to a programming error, five participants in the mismatched condition were presented with one word in the final block twice, while five participants in the matched condition were not presented with this word at all. Two of these ten participants were excluded for being left handed, for the others, this was handled as follows. In the mismatched condition, the eye tracking data for the second presentation of the word were discarded, and the recall score for the word was replaced with the average score in this condition to eliminate any advantages from the double presentation. In the matched condition, the eye tracking data in block 3 were based on four words instead of five (as these participants were only presented with four words in block 3), and the recall score for the fifth word (which was missing) was replaced with the average score in the matched condition (see [6]).

For the eye tracking analyses, we first checked the accuracy of calibration. We had to exclude six participants (all from the matched condition) because of inaccurate calibration (i.e., deviation from the four validation points exceeded 1° in one or more word blocks), leaving 77 participants. We then checked the tracking ratio (i.e., the percentage of time for which the eye tracker actually measured the eye movements) for each trial, and trials were excluded when this ratio was more than two standard deviations below the mean. As a result, the data of two other participants in the matched condition were excluded from further analyses as most of their trials (i.e., 13 and 14 out of the 15 trials) had a poor tracking ratio, leaving 75 participants for the eye tracking analyses (33 in the matched and 42 in the mismatched condition). Eighteen individual trials (six in the matched and twelve in the mismatched condition), divided over 11 participants, were excluded, due to low tracking ratios in these respective trials. Taken together (i.e., including the participants whose trials were entirely excluded), 48 of 1155 trials (4.15%) were excluded due to too low tracking ratios (see [6]). For the remaining 75 participants, mean calibration accuracy for block 1, 2, and 3 was 0.40° (*SD* = 0.12°), 0.40° (*SD* = 0.13°), and 0.40° (*SD* = 0.10°), respectively. Average tracking ratio based on the remaining 1107 trials was 88.82% (*SD* = 10.75%).

For the eye tracking analyses we defined fixations using a 40°/s velocity threshold and a minimal duration of 100 ms (cf. Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka, & Van de Weijer, 2011). We created two areas of interest (AoIs), one for the word (437 x 184 pixels) and one for the picture (536 x 442 pixels). The part of the screen not covered by either word or picture AoI was labelled as 'white space'. We calculated the percentage of fixation time on the Word and Picture AoIs by dividing the total fixation duration (i.e., the sum of fixation duration on the Word AoI, Picture AoI, and white space) by the fixation duration on the Word or Picture AoI respectively. Finally, to explore whether the time spent looking at the mismatching pictures was indeed associated with lower test performance, we computed the correlation between the test performance and fixation duration on the Picture AoI for each condition, per block.

Results

Unless otherwise specified, all data were analysed with a repeated-measures ANOVA with Condition (matched or mismatched) as a between-subjects factor and Word block (first, second, third) as a within-subjects factor.

Test performance

Table 3 shows the scores on the cued recall test per condition per block. Numerically, the mismatched condition performed worse than the matched condition, but there was no statistically significant main effect of Condition, $F(1, 81) = 3.30$, $p = .073$, $\eta_p^2 = .04$. There was a main effect of Word block, $F(2, 162) = 13.06$, $p < .001$, $\eta_p^2 = .14$. Repeated contrasts showed that for all participants, recall performance improved from block 1 to block 2, $F(1, 81) = 9.67$, $p = .003$, $\eta_p^2 = .11$, but not from block 2 to 3, $F(1, 81) = 3.62$, $p = .061$, $\eta_p^2 = .04$ (but see [6]). However, we did not find an interaction effect, $F(2, 162) = 0.23$, $p = .847$, $\eta_p^2 < .01$, probably because –in contrast to Experiment 2- performance of participants in the matched condition also improved with task experience (see Table 3).

Table 3. *Mean (and SD) Recall Performance (max. = 5) and Mental Effort Rating (max. = 9) as a Function of Condition and Word Block in Experiment 3.*

	Performance			Mental effort		
	Block 1	Block 2	Block 3	Block 1	Block 2	Block 3
Matched	2.56 (1.27)	3.04 (1.18)	3.45 (1.37)	5.95 (1.38)	6.69 (1.15)	6.72 (1.32)
Mismatched	2.12 (1.28)	2.82 (1.39)	3.04 (1.61)	6.27 (1.30)	6.55 (0.98)	6.45 (1.36)
Overall	2.31 (1.29)	2.89 (1.31)	3.23 (1.51)	6.12 (1.34)	6.61 (1.06)	6.58 (1.34)

Mental effort

Table 3 shows the descriptive statistics of the mental effort scores in Experiment 3. The analysis revealed no significant main effect of Condition, $F(1, 81) = 0.02$, $p = .894$, $\eta_p^2 < .01$. There was a main effect of Word block, $F(1.63, 131.76) = 7.39$, $p = .002$, $\eta_p^2 = .08$, but no interaction effect, $F(1.63, 131.76) = 2.20$, $p = .113$, $\eta_p^2 = .03$. To follow-up on the main effect of Word block, repeated contrasts were conducted, which showed that, as in Experiment 2, invested mental effort increased from block 1 to block 2, $F(1, 81) = 12.52$, $p = .001$, $\eta_p^2 = .13$, and remained stable from block 2 to block 3, $F(1, 81) = 0.08$, $p = .780$, $\eta_p^2 < .01$.

Eye movement data

As stated earlier, participants were presented with each definition twice, first without and then with a picture present. We analyzed the eye movement data for the second part of the trials, in which a picture was present. As we only postulated hypotheses about the Picture AoI, we did not analyze the data on the word AoI, but for completeness, we do provide the descriptive statistics in Table 4.

Table 4. *Mean (and SD) Percentage of Fixation Time on the Picture and Word AoI as a Function of Word Block and Condition.*

		Block 1	Block 2	Block 3
Picture	Matched	76.79 (13.68)	67.34 (19.87)	59.97 (21.59)
	Mismatched	56.00 (26.24)	35.73 (28.30)	27.58 (24.52)
Word	Matched	22.28 (13.56)	30.93 (19.49)	37.77 (21.31)
	Mismatched	42.14 (26.30)	59.15 (28.31)	69.32 (26.88)

The analysis of the percentage of fixation time on the picture revealed a main effect of Condition, $F(1, 73) = 41.80$, $p < .001$, $\eta_p^2 = .36$, showing that participants in the mismatched condition fixated less on the pictures than participants in the matched condition. Furthermore, we found a main effect of Word block, $F(1.83, 133.72) = 32.22$, $p < .001$, $\eta_p^2 = .31$, indicating that participants allocated less attention to the picture AoI over the course of the experiment. The interaction effect did not reach statistical significance, $F(1.83, 133.72) = 2.93$, $p = .062$, $\eta_p^2 = .04$. However, the pattern of the mean fixation times across conditions seems to suggest that the decrease in fixation time on the picture AoI was larger in the mismatched condition (and see [6]). The descriptive statistics for the word AoI suggest that the attention that was no longer allocated to the picture was now dedicated to the word.

Exploratory analysis: Correlation between picture fixation and test performance. To explore whether more attention to the mismatching pictures was indeed associated with lower recall performance, we computed the correlation between picture fixation time and test performance for each condition, per block. In the matched condition, we found no significant correlations between picture fixation time and test performance in any of the blocks: Block 1, $r(31) = -.104$, $p = .566$; block 2, $r(31) = .031$, $p = .866$; block 3, $r(31) = -.174$, $p = .333$. In the mismatched condition, there was a negative correlation between picture fixation time and test performance, which became stronger with increasing task experience: Block 1, $r(40) = -.120$, $p = .448$; block 2, $r(40) = -.310$, $p = .046$; block 3, $r(40) = -.523$ $p < .001$.

Exploratory analysis: Picture fixation during/after the definition was spoken. The pattern of results in the main analysis, indicating that fixation time on the picture AoI decreased over time, and seemed to decrease more strongly in the mismatched condition, is in line with our hypothesis that participants allocate less attention to the mismatching pictures with increasing task experience. However, due to our experimental set-up (i.e., all trials lasted 11 seconds, while the audio was shorter, $M_{length} = 4.81s$, $SD = 0.88s$.), it is unclear to what extent the picture is being ignored during encoding of the verbal definition. Therefore, an exploratory analysis was performed, analysing the fixation data during the time the audio was playing, and after the audio had stopped (see Appendix A for audio length per Vimmi word definition). Because audio duration differed among words (with some definitions being longer than others), we could not compare absolute fixation times during or after the audio. Instead, we calculated the relative fixation duration by dividing the fixation duration on each AoI by the audio length (for during the audio) or non-audio length (for after the audio ended). Because we assumed that the pictures

---

[6] Exclusion of the eight participants who were affected by the programming error would have a minor impact on our results. Performance scores: The difference between block 2 and block 3 is now significant, $F(1, 73) = 6.26$, $p = .015$, $\eta_p^2 = .08$; eye movement data: The interaction between Condition and Word block is now significant, $F(2, 130) = 4.10$, $p = .019$, $\eta_p^2 = .06$

The exclusion of the 48 trials for which the tracking ratio was to low also had a minor impact on our results for the eye movement data: Without the exclusion, the Condition * Word block interaction was significant, $F(1.82, 136.27) = 3.23$, $p = .042$, $\eta_p^2 = 0.04$.

would hinder the encoding of the spoken definitions, attention allocation *during* the audio was of most interest.

We performed a 3 (Word block: first, second, third) x 2 (Audio: during or after) x 2 (Condition: matched or mismatched) ANOVA on picture fixation duration. The analysis revealed a main effect of Condition, $F(1, 73) = 39.45$, $p < .001$, $\eta_p^2 = .36$, showing that participants in the mismatched condition looked less at the picture AoI than participants in the matched condition (see Table 5). Furthermore, we found a main effect of Word block, $F(2, 146) = 40.50$, $p < .001$, $\eta_p^2 = .36$, indicating that the relative fixation duration on the picture AoI decreased with increasing task experience. This main effect was qualified by an interaction effect between Word block and Condition, $F(2, 146) = 3.87$, $p = .023$, $\eta_p^2 = .05$, showing that the decrease in relative fixation duration was strongest in the mismatched condition. The main effect of Audio was also significant, $F(1, 73) = 374.06$, $p < .001$, $\eta_p^2 = .84$, indicating that after the audio ended, relative fixation duration on the picture AoI decreased. The Audio * Condition and Word block * Audio interaction effects were not significant, respectively, $F(1, 73) = 0.85$, $p = .360$, $\eta_p^2 = .01$, and $F(2, 146) = 2.29$, $p = .105$, $\eta_p^2 = .03$. However, the Word block * Audio * Condition interaction effect was significant, $F(2, 146) = 3.73$, $p = .026$, $\eta_p^2 = .05$. This suggests that the relative picture fixation difference between the matched and mismatched condition decreased more strongly over time *during* the audio than *after* the audio.

Table 5. *Mean (and SD) Relative Fixation Duration on the Picture and Word AoI as a Function of Word Block and Condition, Split out for During Audio and After Audio.*

		Block 1		Block 2		Block 3	
		During	After	During	After	During	After
Picture	Matched	0.70 (0.09)	0.47 (0.16)	0.62 (0.14)	0.40 (0.17)	0.58 (0.17)	0.34 (0.18)
	Mismatched	0.58 (0.17)	0.32 (0.21)	0.40 (0.25)	0.19 (0.20)	0.30 (0.21)	0.14 (0.15)
Word	Matched	0.11 (0.08)	0.22 (0.14)	0.16 (0.11)	0.27 (0.17)	0.19 (0.14)	0.32 (0.17)
	Mismatched	0.22 (0.18)	0.32 (0.17)	0.36 (0.24)	0.46 (0.21)	0.46 (0.25)	0.51 (0.20)

Discussion

The eye movement data of Experiment 3 are in line with our hypothesis that with increasing task experience, participants in the mismatched condition allocated less attention to the pictures. Interestingly, they seemed to pay less attention to the pictures overall (i.e. from the start) than participants in the matched condition. Furthermore, participants in the matched condition also looked less at the pictures with increasing task experience. However, this decrease was stronger in the mismatched condition (although the interaction in the main analysis did not reach statistical significance, but see [6]). Moreover, only in the mismatched condition did we find an increasingly negative correlation between fixation time on the picture AoI and test performance: In blocks 2 and 3, less looking at the mismatching pictures was associated with higher test performance. Thus, it seems that students are indeed capable of adapting their study strategy and start to ignore extraneous information that is not relevant for –and may even interfere with- the learning task.

Similar to Experiment 2, invested mental effort increased from block 1 to block 2 and remained stable from block 2 to block 3. Again, this is likely a result of participants' experience with the first recall test. In contrast to Experiment 2, however, the increased effort investment seems to have translated into better test performance in blocks 2 and 3 in *both* conditions. Whereas we had expected this for participants in the mismatched condition, we did not expect recall performance of participants in the matched condition to improve based on Experiment 2. It seems that participants in Experiment 3 scored somewhat lower on the cued recall test after block 1 than participants in Experiment 2 (Tables 2 and 3), leaving more room for improvement in the matched condition in Experiment 3. We cannot rule out that awareness of being eye tracked had something to do with this lower performance in the first block, although very few studies have addressed effects of eye tracking on viewing behaviour (see Nasiopoulos, Risko, Foulsham, & Kingstone, 2015; Risko & Kingstone, 2011), and we do not know of studies addressing effects on learning or performance, so this remains speculative. It should also be kept in mind that this might just be chance variation due to the smaller sample size in Experiment 3.

The exploratory analysis on viewing behaviour during and after the spoken definition suggests that attending to the mismatching pictures negatively affects word learning in the first block by disrupting the encoding of the verbal definition. That is, overall, more attention was paid to the pictures while listening to the verbal definition than after, but the difference in attention to the pictures decreased more strongly over time in the mismatched than matched condition while hearing the verbal definition than after the audio had ended. In other words, the increase in recall performance in the mismatched condition over time seems to result mainly from being able to ignore the pictures while listening to the verbal definition.

General Discussion

According to the coherence principle in multimedia learning, presenting extraneous information that is not relevant for the learning task should be avoided, because it hinders learning (Mayer & Fiorella, 2014). However, based on eye-tracking research (e.g. Canham & Hegarty, 2010; Haider & Frensch, 1999), we hypothesized that task experience might be a boundary condition for the negative effect of extraneous information on learning. With increasing task experience, learners might adapt their study strategy and ignore extraneous information, which would reduce or lift the negative effects on learning. We assessed this hypothesis in a series of three experiments. Although we did not find evidence for a coherence effect, we did establish that being presented with pictures that mismatched the action words to be learned, had a negative effect on learning outcomes compared to being presented with pictures that matched the action to be learned (Experiment 1). We then confirmed that task experience nullified this negative effect (Experiment 2), and finally, established that participants indeed adapted their study strategy, allocating less attention to the pictures over time, especially to the mismatching pictures (Experiment 3).

Theoretical Relevance

Although we did not find a multimedia effect of matching pictures, or a coherence effect of mismatching pictures, in the sense that neither picture condition differed significantly from the control condition, we did find a negative effect of mismatching compared to matching pictures on word learning. This difference was interesting for our purpose of investigating whether task experience affects processing of mismatching pictures and thereby, learning outcomes. Our findings – as summarized at the beginning of the General Discussion- are relevant for theories of (multimedia) learning and instructional design.

First, they confirm directly (using eye tracking) the mechanism through which the presentation of extraneous information that is irrelevant for and conflicts with the task at hand initially hinders learning. That is, our findings show that such information initially captures attention (i.e. is being processed). This presumably hinders learning by drawing on valuable working memory resources that can no longer be used for processes relevant for learning (Kalyuga & Sweller, 2014, Sweller et al., 2011) and/or by actively conflicting with the to be learned information (cf. Mayer et al., 2007). The additional exploratory analyses of the eye tracking data provide some more insight into the potential mechanisms through which this occurs. First, the reduction in attention to the extraneous information (i.e., mismatching pictures) was associated with an increase in test performance (i.e. negative correlation that became stronger over time). Second, participants seemed to suppress attention to the extraneous information over time particularly while the verbal definition was spoken (i.e. during encoding). Since the reduction in attention was associated with improved performance, this suggests that the extraneous information particularly interferes with encoding. Such direct tests of assumptions about the (attentional) mechanisms underlying multimedia principles are important, because they can support existing ideas about how and why effects on learning occur, or may generate new insights and explanations for these principles (Van Gog & Scheiter, 2010). Thus far, eye-tracking research on the coherence principle is scarce and focussed mostly on effects of presenting 'seductive details' (e.g., Lehman, Schraw, McCrudden, & Hartley, 2007; Rey, 2014; Sanchez & Wiley, 2006). Seductive details are additional but irrelevant pieces of information (pictures, text) that are usually added in an attempt to increase learners' interest or motivation, but have the (unintended) side effect that they often hamper learning. Our findings are in line with those from the eye tracking studies on seductive-details. These studies showed that seductive details hampered test performance by attracting attention, reducing the time learners spent on the relevant learning materials (Lehman et al., 2007). Furthermore, this effect was stronger for participants with lower attention control (Rey, 2014) or working memory capacity (Sanchez & Wiley, 2006).

Second, our findings not only reinforce but also extend prior multi media research by showing that initial negative effects of extraneous information on learning disappear because people learn to adapt their study strategy. Furthermore, this finding also extends prior eye tracking research. It had been shown that people learn to ignore task-irrelevant information during task *performance* as a result of task experience (Haider & Frensch, 1999) or increased prior knowledge (Canham & Hegarty, 2010; Hegarty et al., 2010). Our findings demonstrate that people can also learn to ignore irrelevant information during *learning*, and that this is associated with better learning outcomes. Thus, our findings suggest that task experience may be a boundary condition for the negative effect of extraneous information on learning, because participants stop allocating attention to this information.

Third, the finding that learners seem able to adapt their study strategy to cope with the extraneous information (at least when this information always appears in the same location), without explicit instruction to do so, demonstrates the importance of studying multimedia learning principles over time, taking into account potential changes in study strategy. This focus on spontaneous adaptations of study strategies in multimedia learning is scarce, but fits well with a relatively new line of research in which participants are successfully instructed to self-manage their cognitive load during multimedia learning by changing their study strategy (e.g., Agostinho, Tindall-Ford, & Roodenrys, 2013; Gordon, Tindall-Ford, Agostinho, & Paas, 2016; Roodenrys, Agostinho, Roodenrys, & Chander, 2012), and with research on training multimedia learning strategies (e.g., Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Mason, Pluchino, & Tornatora, 2016; Stalbovs, Scheiter, & Gerjets, 2015). These studies on strategy training, combined with those of the present study, show the importance of studying multimedia learning principles over time, as participants might be able to overcome suboptimal instructional design, with or without explicit instruction to do so.

Although these findings on effects of task experience may bring to mind an “expertise reversal effect” (Kalyuga, 2014), they are rather different. The expertise reversal effect states that learning materials that are essential and non-redundant for novices, become redundant when learners gain or have more prior knowledge, at which point they will no longer aid, and might even hinder learning. Thus, an expertise reversal effect would imply that information becomes extraneous and *starts* to hamper learning as task experience increases, whereas in our study the extraneous information *stops* hampering learning as task experience increases. Moreover, in contrast to the expertise reversal effect, participants in our study did not gain experience with (or knowledge of) the task *content* (i.e., the word definitions) over time, but with the task *presentation*.

Limitations and Future Research

A limitation of the present study lies in the materials used. In Experiments 1a and 1b we failed to establish a coherence effect, as there was no negative effect of mismatching pictures compared to a no picture control condition. We did, however, find a small but significant and consistent negative effect on learning of mismatching compared to matching pictures, which was suitable for addressing our main hypothesis that such a negative effect might diminish with task experience. Furthermore, the learning materials used in the present study were specifically designed to test our hypothesis and as such have rather low ecological validity. Yet, pictures that are irrelevant for the learning task, and might even conflict with it, are ubiquitous in textbooks and e-learning materials, and our study provides a first indication that students can learn to ignore this extraneous information, and that when they do, it no longer hampers their learning.

Future research should investigate the process by which this occurs in more detail. For instance, it is possible that it was relatively easy to suppress attention to the pictures in the present study, because they always appeared in the same location. An interesting question, therefore, is whether participants have truly learned that the content mismatched the definitions. In this case, they would continue to ignore the pictures even when the location would suddenly change (which would indicate a relatively conscious process of attention inhibition).

Another potential limitation of our materials is that our mismatching pictures were not only irrelevant for learning the action word definition, but might also have actively interfered with learning, by depicting another action. Therefore, it would be interesting for future research to further disentangle effects of various types of extraneous information. For instance, one could use decorative pictures that are superficially related to the action, but do not conflict with the to be learned definition (e.g., a picture of a toolbox with a wrench, hammer, and sanding paper inside when learning that 'ifra' means 'to polish or scrape with sandpaper'). An interesting question is whether it would be more difficult for participants to learn to ignore such pictures, as the lack of conflict might make it less obvious that the pictures can/should be ignored.

Moreover, it would be interesting to examine whether our findings would generalize to more ecologically valid and more complex materials that test meaningful learning instead of rote memory (as the present materials required). This would also allow for determining whether our findings would also extend to other types of coherence effects, such as the presentation of unnecessary elaborations or details when learning from illustrated texts. When the learning materials and the type of learning that is required become more complex, it may become less obvious what information is and is not relevant for learning, which may make it harder for learners to adapt their study strategy. Relatedly, the way in which extraneous information is presented may also affect how easily students can learn to ignore it. For instance, it may be harder for learners to start to ignore irrelevant text when it is integrated in pictures (cf. Bobis et al., 1993; Chandler & Sweller, 1991) or to ignore incoherent additional explanations about mechanical systems that interfere with learning the relevant materials (Mayer et al., 2007) than it is to ignore a picture (that always appears in the same location) in its entirety. At the same time, it is arguably even more important to be able to ignore extraneous information in complex materials such as illustrated texts, as such materials already pose a high demand on limited working memory resources.

Finally, it would also be relevant to address the role of individual differences in attention control or working memory (cf. Rey, 2014; Sanchez & Wiley, 2006) in future research. Such individual differences, might affect whether people learn to ignore task-irrelevant information or the rate with which they learn to do so.

Practical Relevance

Despite these limitations, our findings may prove relevant for educational practice. Although our findings require replication with more complex learning materials and more complex types of learning, our study provides a first indication that students can learn to ignore extraneous information, and that when they do, it no longer hampers their learning. Moreover, although learning was 'merely' defined in terms of rote memory in the present study, this also has its place in educational practice. When learning a new language, for example, it is important to first acquire a sufficient vocabulary before learning grammar. As Wilkins (1972, pp. 111-112) put it: "…while without grammar very little can be conveyed, without vocabulary nothing can be conveyed".

Knowing that learners might be able to adapt their study strategies and ignore extraneous information with task experience, is relevant for instructional designers. That is, it is very hard for instructional designers to take into account all multimedia principles at once, because individual learner characteristics may interact with some of the principles. For example, the split attention principle states that information from two mutually referring sources (e.g. text and picture) should be integrated rather than presented separately (Ayres & Sweller, 2014). However, whereas the integrated text may be crucial for novices' understanding, it can become redundant, and start to hamper learning for more advanced learners (Kalyuga & Sweller, 2014). Therefore, it is important to investigate whether and to what extent students themselves are able to adapt their study strategy spontaneously or after training, and we took a first step in that direction.

Conclusion

Concluding, this study suggests that task experience may be a boundary condition for the negative effect of extraneous information on learning, because experience allows learners to change their study strategies to cope with (i.e. ignore) information that interferes with their learning. Future research should establish whether this boundary condition generalizes to more complex learning and other types of extraneous information. If so, this is relevant knowledge for multimedia learning theories as well as for instructional designers

Appendix A

Vimmi Words Used in Experiment 2 and 3 and Audio Duration of the definitions for Experiment 3.

	Vimmi word	Definition of	Duration (s)
1	Repo	to write	4.14
2	Lapo	to iron	4.56
3	Rifa	to saw	6.54
4	Dawu	to inject	4.52
5	Dupi	to hammer	5.28
6	Kune	to paint	3.22
7	Luko	to scrub	4.56
8	Bepa	to screw	5.88
9	Redu	to stir	6.07
10	Buto	to cut	4.82
11	Lozu	to pour	3.89
12	Tari	to erase	5.22
13	Kori	to stamp	3.79
14	Lefa	to beat	5.24
15	Ifra	to sand	4.43

Chapter 3

With Task Experience Students Learn to Ignore the Content, not Just the Location of Irrelevant Information

This chapter has been published as:

Rop, G., Verkoeijen, P. P. J. L., & Van Gog, T. (2017). With task experience students learn to ignore the content, not just the location of irrelevant information. *Journal of Cognitive Psychology, 29*, 599-606. doi:10.1080/20445911.2017.1299154

Abstract

Presentation of irrelevant additional information hampers learning. However, using a word-learning task, recent research demonstrated that an initial negative effect of mismatching pictures on learning no longer occurred once learners gained task experience. It is unclear, however, whether learners consciously suppressed attention to the content of the mismatching pictures. Therefore, we examined the effects of a picture location change towards the end of the learning phase: for half of the participants, the picture location was changed after they gained task experience. If participants only ignore the *location* of mismatching pictures, word learning in the mismatched condition should be hampered after the location change. Changing the location of the mismatching pictures did not affect recall in the mismatched condition, but, surprisingly, the location change did hamper learning in the matched condition. In sum, it seems that participants learned to ignore the *content*, and not just the *location* of the irrelevant information.

Introduction

The ”multimedia effect” indicates that learning improves when study tasks or materials combine pictorial and verbal representations of the content (Butcher, 2014). However, this beneficial effect on learning only occurs when both representations are crucial for understanding the subject at hand. When one source of information is extraneous, that is, not relevant for learning, it will hinder learning (Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). For example, learning is hampered when interesting information is added to enrich materials (i.e., seductive details, e.g., Harp & Mayer, 1998); when learning materials are unnecessarily elaborate, presenting textual explanations with self-explanatory diagrams (e.g., Chandler & Sweller, 1991), or providing details and examples whereas a concise summary would suffice (e.g., Mayer, Bove, Bryman, Mars, & Tapangco, 1996); or when information on related systems is presented when learning about a specific system (Mayer, DeLeeuw, & Ayres, 2007).

The negative effects of extraneous information on learning arise because learners attend to, process, and attempt to integrate the extraneous information with the essential information, which unnecessarily depletes working memory resources required for learning (Mayer, 2014; Sweller, Ayres, & Kalyuga, 2011). Moreover, in some cases, the content of the additionally presented information may actively interfere with learning the essential information (e.g., Mayer et al., 2007). However, eye-tracking studies have shown that participants learn to ignore extraneous information with task experience (Haider & Frensch, 1999) or explicit instruction (Hegarty, Canham, & Fabrikant, 2010). Therefore, task experience might be a boundary condition to the negative effect of extraneous information on learning: If people learn to ignore such information with task experience, it should no longer hamper learning.

A recent study yielded evidence in line with this hypothesis (Rop, Van Wermeskerken, De Nooijer, Verkoeijen, & Van Gog, in press). Participants learned the definitions of fifteen words (from an artificial language called Vimmi; see Macedonia & Knösche, 2011) in three blocks of five words, with a recall test after each block. After the first block, recall performance was lower when words were coupled with mismatching pictures than with matching pictures; however, once participants had some experience with the task (i.e., in blocks 2 and 3), the mismatching pictures no longer hampered recall performance compared to the matching pictures (Experiment 2). A follow-up experiment, employing eye-tracking methodology to study learners' attention allocation, showed that learners adapted their study strategy with increasing task experience and started to ignore the mismatching pictures more strongly than the matching pictures.

Because the mismatching pictures always appeared at a fixed location, it is an open question whether learners consciously suppressed attention to the pictures because they were aware that the content was irrelevant for the task at hand. One way to answer this question is by systematically changing the location of the pictures for half of the participants after they have accumulated task experience (i.e., in the third block of words; see Figure 1 for an impression of the location change). If they learned to suppress attention to the *location*, learning should be negatively affected in the mismatched condition with a location change (because the change reinstates attention to the pictures, at least briefly) compared to all other conditions. However, if participants learned that the *content* is irrelevant, they would be expected to actively suppress attention to the pictures regardless of the location and performance should not be lower in the mismatched condition with a location change compared to all other conditions.

Another possibility is that a location change will only briefly hamper learning. This hypothesis is based on the *signal-suppression hypothesis* (Gaspelin, Leonard, & Luck, 2015; Sawaki & Luck, 2010), which states that a combination of bottom-up and top-down influences determines attention paid to a stimulus. While a location change might briefly attract attention due to saliency of a stimuli unexpectedly appearing at a different location (bottom-up attention influence, cf. Remington, Johnston, & Yantis, 1992), awareness that the stimulus does not contain useful content (top-down attention influence) would suppress attention to the picture. Consequently, in our learning task, the location change of the mismatching pictures might only hamper learning for the first few words.

Present Experiment

In the present experiment, participants learned fifteen word definitions in three blocks of five words, with either matching (depicting the action to be learned) or mismatching (depicting another action) pictures added. In two conditions, the pictures were presented underneath the word during the whole experiment (these conditions replicate the conditions in Rop et al., in press), while in the other two conditions the location of the pictures changed in block 3, in which they were now presented above the word.

We hypothesized that if learners are aware that the mismatching pictures are irrelevant for their learning, they would suppress attention to the pictures even after the location changes, in which case the change would not influence word learning (either in block 3 as a whole or for the first few words) compared to all other conditions. If they only ignored the location, however, recall performance in the mismatched condition should be negatively affected after the location change (at least for the first few words in block 3). We also performed a direct replication experiment (Experiment 1b) as one finding from Experiment 1a was interesting but surprising.

Method

Participants and Design

Participants (Experiment 1a: n = 429, Experiment 1b: n = 485) were recruited on Amazon's Mechanical Turk (Buhrmester, Kwang, & Gosling, 2011) and were paid 1.50 US dollar for their participation. A-priori defined post-hoc exclusion criteria were: Being left handed (n = 67, n = 80); being a non-native English speaker (n = 11, n = 4); participating in a noisy environment (i.e., a self-reported score of 7 or higher on a 9 point scale, n = 5, n = 5); and taking notes during the learning phase (n = 8, n = 10). Furthermore, some participants were excluded for misunderstanding the instructions (i.e., they wrote down the names of the pictures instead of the word definitions which they were instructed to learn; n = 8, n = 8); and some participants were excluded because they encountered technical difficulties (n = 2, n = 4). Finally, one participant in Experiment 1a did not have an MTurk ID and was excluded, while in Experiment 1b we excluded all participants that already participated in Experiment 1a (n = 22).

This left 327 participants in Experiment 1a (M_{age} = 37.50 years, SD = 11.71 years, range 18-68; 199 females), who were randomly distributed over four conditions resulting from a 2 x 2 design with between-subjects factors "Picture Match" (matching vs. mismatching) and "Location Change" (yes vs. no): matching pictures no location change (matched condition, n = 72), matching pictures with location change (matched-change condition, n = 90), mismatching pictures no location change (mismatched condition, n = 87), and mismatching pictures with location change (mismatched-change condition, n = 78). In Experiment 1b, 352 participants were left (M_{age} = 36.25 years, SD = 11.02 years, range 18-71; 180 females), who were randomly distributed over the matched (n = 91), matched-change (n = 89), mismatched (n = 86), and mismatched-change (n = 86) conditions.

Materials and Procedure

The learning materials were programmed in Qualtrics software (Qualtrics, Provo, UT). Participants learned the definitions of fifteen Vimmi words in three blocks of five words, with a recall test after each block. Each word was coupled to the definition of an action verb (e.g., “ifra” means “to polish or scrape with sandpaper”). Participants saw the word printed on screen and heard the spoken definition of the word they had to learn twice (each presentation lasted 11 seconds and the program automatically progressed). A matching or a mismatching picture accompanied the word the second time participants heard the definition. In the two conditions without a location change, the picture was always presented underneath the word. In the two change conditions, the picture was presented underneath the word in block 1 and 2, but above the word in block 3 (see Figure 1).

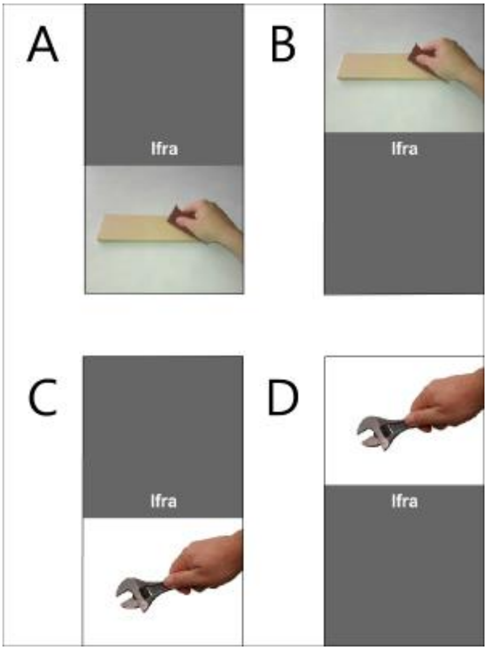


Figure 1. Example materials. The spoken definition (e.g., Ifra means to polish or scrape with sandpaper) is presented twice, the second time accompanied by a picture. In the matched (1A) and mismatched (1C) conditions, this picture was always presented underneath the word. In the matched-change (1B) and mismatched-change (1D) conditions the picture was presented underneath the word in block 1 and 2, but the location of the picture changed in block 3, in which the picture was now presented above the word.

Participants' knowledge of the definition was tested with a cued recall retention test after each block, in which they were presented with the written word and had to type in the associated definition as literally as possible[1]. A block always consisted of the same 5 words, but the order of the blocks was randomized using a Latin-square design, which resulted in 12 lists used for the experiment. There were no breaks between blocks. The experiment lasted about twenty minutes.

Scoring

Participants were awarded 1 point if they provided a complete definition on the cued recall test (e.g., "to polish or scrape with sandpaper" for the word "ifra"). When part of the definition was missing, they received 0.5 point (e.g., "to polish"). If they did not provide a definition, or if it was completely wrong, 0 points were awarded (e.g., "to remove something written by wiping" which was the definition of another word in that block). So, every participant could score a maximum of five points on each test. A random subset of the data (11.0% in Experiment 1a and 10.2% in Experiment 1b) was scored by a second rater, and interrater reliability was high (κ = .91 in Experiment 1a and κ = .84 in Experiment 1b).

Results

In all analyses, a significance level of .05 was maintained, and when the sphericity assumption was violated, the Greenhouse-Geisser correction is reported. Effect size measures used were partial eta-squared and Cohen's *d*. Both can be interpreted in terms of small ($\eta_p^2 \sim .01$, $d \sim 0.2$), medium ($\eta_p^2 \sim .06$, $d \sim 0.5$), and large ($\eta_p^2 \sim .14$, $d \sim 0.8$) effect sizes (Cohen, 1988). First, to check whether we could replicate prior findings by Rop et al. (in press) that performance is initially hampered by mismatching pictures, we performed a mixed ANOVA on recall performance with Word Block (first or second) as within-subjects factor and Picture Match (matching or mismatching) as between-subjects factor. Then, to test our hypothesis concerning the effects of the location change on recall performance we conducted 2 x 2 ANOVAs with Picture Match (matching or mismatching) and Location Change (yes or no) as

[1] We also explored whether there were differences in experienced cognitive load among the conditions, by asking participants to indicate how much mental effort they invested in learning the words on a nine point rating scale (Paas, 1992), ranging from one (very, very low effort) to nine (very, very high effort). Because of word limits we do not report these data (the only significant finding concerned a main effect of Picture Match in Experiment 1b, $F(1, 348) = 5.96$, $p = .015$, $\eta_p^2 = .02$, indicating that participants in the mismatched condition invested more mental effort).

between-subjects factors on recall performance in block 3. Finally, we investigated effects on word level within block 3 by calculating the recall performance per word in that block and performing two repeated-measures ANOVA's (for the mismatched and matched condition separately), with Serial Position (1, 2, 3, 4, and 5) as within-subjects factor and Location Change (yes or no) as between-subjects factor.

Check: Did Mismatching Pictures Initially Hamper Learning?

Table 1 shows the results on recall performance in block 1 and 2 of Experiment 1a and 1b, and of Experiment 2 of Rop et al. (in press). Both Experiment 1a and 1b showed a significant main effect of Word Block indicating that recall performance improved in both conditions from block 1 to block 2 (1a: $F(1, 325) = 10.76$, $p = .001$, $\eta_p^2 = .03$; 1b: $F(1, 350) = 15.58$, $p < .001$, $\eta_p^2 = .04$). However, we did not replicate the interaction between Word Block and Picture Match that was found in the study by Rop et al. (in press; 1a: $F(1, 325) = 1.17$, $p = .347$, $\eta_p^2 < .01$; 1b: $F(1, 350) = 1.86$, $p = .174$, $\eta_p^2 = .01$). Because the pattern in the data seemed consistent with our hypothesis and the interaction effect was small in the prior study, we decided to analyse the combined data from the prior and current study in order to get an estimate of the combined effect of these three studies. To do so, we performed a mixed ANOVA with Picture Match (matching or mismatching) and Experiment (Rop et al., Experiment 2; Experiment 1a, and Experiment 1b from the present study) as between-subjects factors and Word Block (first or second) as a within-subjects factor. In this analysis, the interaction between Word Block and Picture Match was significant, $F(1, 777) = 6.11$, $p = .014$, $\eta_p^2 = .01$, while the three-way interaction Word block x Picture Match x Experiment was not, $F < 1$. The lack of a three-way interaction suggests that the patterns of results in the experiments are comparable. Therefore we followed-up on the interaction between Word Block and Picture Match with one-tailed *t*-tests. These tests showed that participants in the matched condition had better recall performance than participants in the mismatched condition in block 1, $t(781) = 2.31$, $p = .011$, $d = 0.17$, but not in block 2, $t(781) = 0.06$, $p = .475$, $d < 0.01$.

Table 1. *Mean (and SD) Recall Performance (max. = 5) as a Function of Picture Match and Location Change in Experiment 1a, Experiment 1b, and Rop et al., Experiment 2.*

	Block 1	Block 2	Block 3
Experiment 1a (*n* = 327)			
Matched	3.04 (1.50)	3.17 (1.65)	3.24 (1.63)
Matched-change	2.89 (1.38)	3.17 (1.53)	2.85 (1.54)
Mean	2.96 (1.43)	3.17 (1.58)	
Mismatched	2.71 (1.57)	3.09 (1.53)	3.16 (1.52)
Mismatched-change	3.01 (1.49)	3.40 (1.54)	3.49 (1.46)
Mean	2.86 (1.54)	3.23 (1.54)	
Experiment 1b (*n* = 352)			
Matched	2.58 (1.39)	2.77 (1.48)	2.73 (1.52)
Matched-change	2.63 (1.85)	2.90 (1.85)	2.52 (1.54)
Mean	2.61 (1.63)	2.83 (1.67)	
Mismatched	2.42 (1.41)	2.65 (1.43)	2.73 (1.49)
Mismatched-change	2.23 (1.39)	2.94 (1.45)	3.01 (1.42)
Mean	2.32 (1.40)	2.79 (1.44)	
Rop et al., Exp 2 (*n* = 104)			
Matched	3.15 (1.46)	3.21 (1.51)	3.14 (1.63)
Mismatched	2.51 (1.51)	3.17 (1.45)	3.13 (1.44)

Hypothesis: Is Recall Performance in Block 3 Affected by the Picture Location Change?

Experiment 1a. The recall performance in block 3 is shown in Table 1. There was no main effect of Picture Match, $F(1, 323) = 2.66$, $p = .104$, $\eta_p^2 = .01$, or Location Change, $F < 1$ on recall performance, but we did find a significant interaction between Picture Match and Location Change, $F(1, 323) = 4.61$, $p = .032$, $\eta_p^2 = .01$. Bonferroni corrected follow-up *t*-tests (two-tailed) indicated that, in the absence of a change, recall performance between the matched and mismatched conditions was comparable, $t(157) = 0.35$, $p > .999$, $d = 0.06$, 95% CI for the difference in means = [-0.58; 0.41]. Surprisingly, after a change, recall performance was *higher* in the mismatched condition than in the matched condition, $t(166) = 2.77$, $p = .012$, $d = 0.43$, 95% CI = [0.18; 1.10].

Experiment 1b. Again, there was no main effect of Picture Match, $F(1, 348) = 2.31$, $p = .129$, $\eta_p^2 = .01$, or Location Change, $F < 1$, and—in contrast to Experiment 1a—the interaction effect was not statistically significant, $F(1, 348) = 2.28$, $p = .132$, $\eta_p^2 = .01$, although the pattern of results as well as the effect size was comparable to Experiment 1a. Therefore, we exploratively conducted the same set of Bonferroni corrected follow-up tests as in Experiment 1a. These results were also in the same direction as in Experiment 1a, although not statistically significant. In the absence of a change, recall performance between the two picture conditions was comparable, $t(175) = 0.01$, $p > .999$, $d < 0.01$, 95% CI = [-0.45; 0.45], while recall performance seemed higher in the mismatched condition than in the matched condition after a location change, $t(173) = 2.16$, $p = .066$, $d = 0.33$, 95% CI = [0.04; 0.93].

Combined analysis. We ran a combined analysis of Experiment 1a and 1b[2], as these experiments are a direct replication of each other. We performed a 2 x 2 x 2 ANOVA with Picture Match (matching or mismatching), Location Change (yes or no) and Experiment (Experiment 1a and 1b) as between-subjects factors. This analysis revealed a significant interaction between Picture Match and Location Change, $F(1, 671) = 15.52$, $p = .009$, $\eta_p^2 = .01$, while the three-way interaction Picture Match x Location Change x Experiment was not significant, $F < 1$ (again suggesting that the Experiments are comparable). The follow-up tests showed that, in the absence of a change, recall performance between the two picture conditions did not differ, $t(334) = 0.07$, $p = .994$, $d = 0.01$, 95% CI = [-0.34; 0.32], while recall performance was higher in the mismatched condition than in the matched condition after a location change, $t(341) = 3.39$, $p = .001$, $d = 0.37$, 95% CI = [0.23; 0.87]. This combined analysis gives a better estimation of the true effect of a location change, which is a small-to-medium effect.

Hypothesis: Is Recall Performance in Block 3 Affected on Word Level?

Experiment 1a. Table 2 presents the recall performance data at the word level in block 3. Our main objective of this analysis was to explore whether a negative effect of location change would occur in the first few serial positions of block 3 for mismatching pictures but not for matching pictures. Therefore, we will only report

---

[2] Based on an anonymous Reviewer's suggestion.

on the interaction between Location Change and Serial Position, which was not significant for the matched, $F < 1$ and mismatched condition, $F < 1$.

Experiment 1b: Again, we did not find an interaction between Location Change and Serial Position for both conditions: matched, $F < 1$; mismatched, $F(3.63, 616.67) = 1.21$, $p = .307$, $\eta_p^2 = .01$.

Table 2. *Mean (and SD) Recall Performance on the Words in Block 3 as a Function of Picture Match and Location Change in Experiment 1a and 1b.*

	Experiment 1a				Experiment 1b			
	Matched		Mismatched		Matched		Mismatched	
	No Change	Change	No Change	Change	No Change	Change	No Change	Change
1	.70 (.43)	.66 (.46)	.66 (.46)	.69 (.43)	.63 (.45)	.60 (.45)	.56 (.47)	.71 (.43)
2	.60 (.46)	.52 (.43)	.59 (.47)	.69 (.44)	.49 (.42)	.43 (.42)	.51 (.45)	.55 (.43)
3	.60 (.45)	.51 (.42)	.57 (.43)	.60 (.42)	.47 (.42)	.44 (.42)	.48 (.44)	.48 (.43)
4	.64 (.42)	.53 (.45)	.60 (.43)	.64 (.43)	.48 (.42)	.47 (.41)	.48 (.46)	.56 (.44)
5	.71 (.39)	.63 (.44)	.74 (.38)	.88 (.30)	.67 (.41)	.60 (.46)	.70 (.40)	.71 (.39)
∑	3.24 (1.63)	2.85 (1.54)	3.16 (1.52)	3.49 (1.46)	2.73 (1.52)	2.52 (1.54)	2.73 (1.49)	3.01 (1.42)

Explorative Analysis: Does Recall Performance Within Conditions Change from Block 2 to 3?

To exploratively follow up on the unexpected finding that recall performance was higher in the mismatched that in the matched condition when a change was present (see Table 1), we performed Bonferroni corrected paired *t*-tests to compare the performance in block 2 and 3 in all four conditions of Experiment 1a and 1b. The results of Experiment 1a suggest that recall performance was lower in block 3 compared to block 2 in the matched-change condition, $t(89) = 2.11$, $p = .072$, $d = 0.21$, 95% CI = [-0.02; -0.61], whereas performance remained stable across block 2 and 3 in the other three conditions, minimum $p = .860$, maximum $d = 0.07$. In Experiment 1b, again, there seemed to be a performance drop in the matched-change condition from block 2 to 3, $t(88) = 2.07$, $p = .084$, $d = 0.22$, 95% CI = [-0.01; -0.74], which did not occur in the other conditions, minimum $p < .999$, maximum $d = 0.05$.

Discussion

Prior research has shown that presenting learners with extraneous information that is irrelevant for the task at hand, hampers their learning (Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). However, a recent study comparing the effect of matching and mismatching pictures on word learning, suggested that task experience might be a boundary condition to this effect (Rop et al., in press). The negative effect on learning was present initially but no longer occurred once learners gained task experience, because they started to ignore the irrelevant information. However, the mismatching pictures always appeared at a fixed location. Therefore, it was unclear whether learners *consciously* suppressed attention to the pictures because they were aware that the content was irrelevant for the task at hand. The aim of the present study was to address this question by systematically changing the location of the pictures for half of the participants after they have accumulated task experience. Our results indicated that changing the picture location influenced recall performance, albeit in an unexpected way. The location change in block 3 resulted in poorer recall in the matched condition compared to the mismatched condition, and an explorative follow-up analysis suggested that recall performance decreased in the matched condition from block 2 to block 3, while it remained stable in all other conditions. Note that this analysis was not statistically significant after Bonferroni correction. However, the effect sizes in Experiment 1a and Experiment 1b were almost equal (d = 0.21 and 0.22). Combined, these findings suggest that changing the location of matching pictures seemed to have a small negative effect on word learning.

Eye-tracking data from the study by Rop et al. (in press) showed that matching pictures continuously attracted a substantial amount of attention, from an average of 76% of fixation time in block 1 to 60% in block 3, over the course of the experiment. Thus, in the present study, when the location of these pictures suddenly changed in block 3, participants might have wondered why the location of the pictures changed, which would distract from learning the definitions. This distraction might have hampered learning as participants focused more on the changing picture location, and less on encoding the actual definition. Future research could address the plausibility of this explanation by measuring learners' visual attention allocation using eye-tracking methodology to see whether they anticipated on the picture appearing in the other location, or by interviewing them after the experiment. More importantly for our hypotheses, the fact that the location change did not affect performance in the mismatched condition suggests that students were aware that the content was irrelevant for their learning of the word definitions and that they continued to ignore these pictures.

Limitations and Future Research

A limitation of the present study is that we did not directly measure visual attention allocation, but the performance data suggest that the mismatching pictures must have been consciously ignored via top-down influences, because otherwise a drop in performance compared to block 2 would have occurred. Furthermore, within block 3 we did not find a negative effect of mismatching pictures on the first words, so even if the location change attracted learners' attention initially (i.e., stimulus-driven, bottom-up influences; cf. Remington et al., 1992), it seems to have been suppressed quickly (cf. Gaspelin et al., 2015; Sawaki & Luck, 2010). Possibly, participants were able to ignore the mismatching pictures by redirecting their attention to the artificial language word that was shown on the screen (in the study by Rop et al., in press, attention to the word increased from an average of 42% in block 1 to 69% in block 3 in the mismatched condition).

Another possible limitation of the present study could be that we only replicated the initial finding that mismatching pictures have a negative effect on learning compared to matching pictures when we combined the results of multiple experiments. Note though, that the pattern of means of recall performance was consistent over all experiments: Participants learning with mismatching pictures score lower in block 1 than participants learning with matching pictures. Secondly, because we were able to include multiple experiments in the combined analysis, we had a large sample size, which means that we can be fairly certain that the effect exists, although it is small. Finally, the small effect size is consistent with prior studies using the same materials (Rop et al., in press; De Nooijer, Van Gog, Paas, & Zwaan, 2013) and may perhaps be due to the relatively low complexity of the learning materials. All things considered, we can regard this replication attempt a modest success, although the effect size for the crucial interaction we found is much smaller than the effect size reported in Rop et al. (in press). Future research should address whether these findings would replicate with more complex multimedia learning materials, such as expository texts combined with explanatory pictures. Such materials might induce larger effect sizes, and would provide evidence that task experience is a robust boundary condition to the negative effects of irrelevant information on learning.

Practical Implications

Our results may also be relevant for educational practice. Although the study by Rop et al. (in press) already showed that over time, students are able to adapt their study strategy and ignore irrelevant information, it was an open question whether participants consciously supressed attention to the pictures. The results of the present study suggest that they truly learned to ignore the *content*, and not just the *location* of the irrelevant information. Because information that is relevant for novices might become irrelevant for advanced learners, it is important for instructional designers to know that students seem to be able to adapt their study strategies in multimedia learning. Interestingly, our findings do suggest that instructional designers might want to be careful with changing the location of *relevant* information after learners have gained experience with the task, as our findings suggest that this can have a (small) negative effect on learning. Future research should attempt to replicate these findings in other materials, however, before clear instructional design guidelines can be derived.

Chapter 4

Effects of Task Experience and Layout on Learning from Text and Pictures with or without Unnecessary Picture Descriptions

This chapter has been submitted for publication as:

Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2017). *Effects of Task Experience and Layout on Learning from Text and Pictures with or without Unnecessary Picture Descriptions*. Manuscript submitted for publication.

Abstract

The presentation of extraneous (i.e., irrelevant or unnecessary) information may hamper learning with multimedia. The present study examined whether people can learn to ignore unnecessary information with increasing task experience, and whether this depends on the layout of that information. Participants learned about the process of mitosis from a multimedia slideshow, with each slide presenting a combination of expository text and a picture on one of the stages in the process. Slides either contained no unnecessary text (control condition), or unnecessary text (i.e., merely describing the picture) either integrated in the picture (integrated condition), or presented underneath the picture (separated condition). Experiment 1 showed that the addition of separated unnecessary information hampered learning compared to the control condition, but only when the unnecessary information was presented at the right-hand side of the screen. We could not replicate these results in Experiment 2, in which we also employed eye tracking. The eye movement data did confirm, however, that participants attended less to the unnecessary information with increasing task experience. This finding suggests that students seem to adapt their study strategy and learn to ignore information that is unnecessary for their learning, even if processing this information does not hamper their learning.

Introduction

According to well-known multimedia design principles, presenting extraneous (i.e., irrelevant or unnecessary) information in study material should be avoided, as it hinders learning (for reviews, see Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). A recent study suggested, however, that task experience may be a boundary condition for the negative effect of extraneous information on learning, at least when this information is *pictorial*, and *irrelevant* for the learning task (i.e., Rop, Van Wermeskerken, De Nooijer, Verkoeijen, & Van Gog, in press). In this study on word learning, participants who were presented with pictures that depicted the to-be-learned action word performed better initially (i.e., on the first set of words) than participants who were presented with pictures showing a different action. However, this difference between relevant and irrelevant picture conditions disappeared as learners gained experience with the task (i.e., on later sets of words). Eye-tracking data suggested that the initial negative effect on learning disappeared because learners started to ignore the irrelevant pictures with task experience.

However, because the extraneous information was pictorial and obviously irrelevant (i.e., it mismatched the verbal information that participants had to remember), it is an open question whether task experience would have similar effects when the extraneous information is textual (e.g., a text describing the elements of a picture) and unnecessary rather than irrelevant (i.e., in the sense that the information provided by the text is relevant for the learning task but not necessary as it can also be inferred from the picture). The present study addressed this question.

Effects of Extraneous Information on Learning

While learning from multimedia materials, that is, materials in which text (either spoken or written) and pictures (either static or dynamic) are combined (Mayer, 2014), a learner first has to select the relevant information from the text and picture (by attending to it). Subsequently, this information has to be organised into a coherent cognitive structure in working memory, and has to be integrated with prior knowledge from long-term memory (Mayer, 2014). When either one of these processes (i.e., selection, organisation, or integration) is disrupted, learning is hampered. The presentation of extraneous information in multimedia learning materials may hamper learning when it captures attention, because working memory capacity is limited (e.g., Baddeley, 2000; Cowan, 2001; Miller 1956) and processing this extraneous information that is not conducive to learning reduces the working memory resources available for the selection, organization and integration of information that is essential for learning.

The negative effect of extraneous information processing on learning have been demonstrated with a variety of materials and types of extraneous information. For instance, it has been shown to occur when multimedia learning materials are enriched with interesting and entertaining information (i.e., seductive details; Harp & Mayer, 1998, Lehman, Schraw, McCrudden, & Hartley, 2007; Mayer, Heiser, & Lonn, 2001; Moreno & Mayer, 2000; Rey, 2014; Sanchez & Wiley, 2006), when information on related systems is presented when learning about a specific system (Mayer, DeLeeuw, & Ayres, 2007), or when mismatching pictorial information is provided when learning word definitions (De Nooijer, Van Gog, Paas, & Zwaan, 2013; Hald, Van den Hurk, & Bekkering, 2015; Rop, Van Wermeskerken et al., in press; Rop, Verkoeijen, & Van Gog, 2017). Moreover, the effect has been demonstrated when text accompanying pictures or animation is presented in both spoken and written form (e.g., Craig, Gholson, & Driscoll, 2002, Mayer et al., 2001; but see Mayer & Johnson, 2008; Yue, Bjork, & Bjork, 2013), when self-containing diagrams are accompanied by textual explanations (Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1991), and when unnecessary details and examples are added to learning materials (e.g., Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Reder & Anderson, 1982).

In all these studies, the extraneous information was either irrelevant or unnecessary, depending on its relation with the learning goal. Irrelevant information is unrelated to the learning goal (e.g., seductive details, information about related systems, and mismatching information), while unnecessary information is related to the learning goal, but not necessary for learning because the information is presented twice (e.g., the same spoken and written text accompanying an illustration, self-contained diagrams with unnecessary textual explanations) or is unnecessarily elaborate (e.g., unnecessary details and examples).

As mentioned above, presenting such extraneous information hampers learning because it captures learners' attention, and learners spend valuable cognitive resources on processing this information that is not conducive to learning. However, recently, evidence emerged that people might learn to ignore extraneous information with increasing task experience.

Task Experience

Inspired by eye-tracking studies showing that participants may learn to ignore task-irrelevant information as a consequence of task experience (Haider & Frensch,

1999) or explicit instruction (Canham & Hegarty, 2010; Hegarty, Canham, & Fabrikant, 2010), Rop, Van Wermeskerken et al. (in press) investigated whether people would learn to ignore extraneous information once they have gained some experience with a learning task. If learners would indeed start to ignore extraneous information, it no longer uses up working memory resources, and the negative effect on learning outcomes should therefore become smaller or disappear entirely. The authors tested this hypothesis using a word-learning task. Participants had to learn the definition of an action word from an artificial language (e.g., 'ifra' means 'to polish or scrape with sandpaper') and were shown either relevant pictures that depicted the to-be-learned action or irrelevant pictures showing a different action.

Participants who were presented with the irrelevant pictures initially (i.e., after the first set of words) remembered fewer word definitions than participants who were presented with relevant pictures. However, this difference between relevant and irrelevant picture conditions disappeared as learners gained experience with the task (i.e., on later sets of words). Eye tracking data suggested that the initial negative effect on learning disappeared because learners started to ignore (i.e., allocate less attention to) the irrelevant pictures with task experience (Rop, Van Wermeskerken et al., in press). Moreover, a subsequent study provided evidence that learners started to ignore the pictures based on their content (Rop, Verkoeijen, et al., 2017).

These results suggest that with task experience, learners adapt their study strategy and start to ignore irrelevant information, thereby diminishing the negative effect of irrelevant information on learning. Thus, task experience may be a boundary condition for the negative effect of extraneous information on learning, because participants stop allocating attention to this information. It is important to establish potential boundary conditions as they describe the limits of generalizability of scientific theories (Busse, Kach, & Wagner, 2016; Whetten, 1989). However, as the extraneous information in the study by Rop, Van Wermeskerken et al. (in press) was pictorial, obviously irrelevant (i.e., it mismatched the verbal information that participants had to remember), and not integrated with other information, it is an open question whether task experience would have similar effects when the extraneous information is textual (e.g., a text describing the elements of a picture), unnecessary rather than irrelevant (i.e., in the sense that the information provided by the text is relevant for the learning task but not necessary as it can also be inferred from the picture), and integrated with relevant information (e.g., unnecessary text integrated with a picture).

There are several reasons why task experience might not have a similar effect (i.e. might not help students to learn to ignore extraneous information) under those circumstances. First, textual information may be harder to ignore than pictorial information as learners often focus more quickly and more strongly on text than on the associated pictures (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010). Second, unnecessary information might be harder to ignore than irrelevant information, as it is likely less obvious for learners that unnecessary information is extraneous to their learning process. Finally, whereas extraneous information that is presented separated from the relevant information might be relatively easy for participants to ignore, that might be more difficult when it is integrated with relevant information. The present study addressed these questions.

The Present Study

The present study aimed to answer two questions: 1) Do students learn to ignore unnecessary textual information with increasing task experience, and 2) is the unnecessary textual information more difficult to ignore when it is integrated with relevant information? We conducted two experiments in which participants learned about the process of mitosis using a multimedia slideshow. The slides consisted of a text explaining the process of mitosis (relevant text), and a picture of the visuo-spatial appearance of the cell in that particular stage of mitosis. In two conditions, a description of the picture components (unnecessary text) was added to the slide, either separated from (separated condition), or integrated in (integrated condition) the pictorial information. This information was relevant for the learning goals, but it was unnecessary as it provided a textual description of the picture, while a picture is generally a better representation of visuo-spatial content (Levie & Lentz, 1982; Schmidt-Weigand & Scheiter, 2011). In the third condition, only the relevant text and the picture were presented on the slide (control condition)

Experiment 1 investigated, by measuring learning immediately after each slide, whether an initial negative effect of unnecessary information would occur; whether this negative effect would decrease (or even disappear) as participants gained task experience; and whether this decrease would be stronger when the unnecessary text was presented separated from the picture (i.e., separated unnecessary text would be easier to ignore than integrated unnecessary text). Because we expected that the processing of unnecessary information would increase cognitive load, participants were asked to rate how much mental effort they invested in learning the materials immediately after each slide (as an indicator of how much cognitive load participants experienced: Paas, Tuovinen, Tabbers, & Van Gerven, 2003). We also asked participants to rate how much mental effort they invested during the test phase after each slide, as participants who gained more knowledge during the learning phase should be able to attain higher test performance with less investment of mental effort (Van Gog & Paas, 2008). We expected that participants in the unnecessary-information conditions would initially invest more mental effort during the learning and test phase than participants in the control condition, while this difference should decrease (or even disappear) as participants gained task experience (at least in the separated condition). Experiment 2 was a direct replication of Experiment 1, apart from the fact that eye tracking was employed to directly study attention allocation processes.

Experiment 1

Method

Participants and design

Initially, 96 individuals participated in the study, recruited via the university's online recruitment systems. Due to an error in one of those systems, it turned out that two participants had already graduated and therefore they were excluded from the sample. The final sample comprised 94 undergraduate students from a Dutch University (M_{age} = 21.74 years, *SD* = 2.55 years; 65 female) who participated for course credit or a small fee of 5 euro. They were randomly assigned to one of the three conditions: control (*n* = 32), integrated (*n* = 31), and separated (*n* = 31).

Materials

The materials were designed and presented using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA).

Prior knowledge test. Participants' prior knowledge was tested with four multiple-choice questions about the process of mitosis (e.g., what is mitosis?) with four possible answers (e.g., the correct alternative: "A process in which the nucleus and duplicated chromosomes of a cell divide and are evenly distributed"). During this test, and all other tests used in this experiment, participants had to guess when they did not know an answer, as the program would not progress unless they answered a question.

Learning materials. The learning materials consisted of a slideshow in which six slides described and depicted the process of mitosis. During mitosis, the nucleus and chromosomes of a cell are duplicated and are divided over 2 new daughter cells.

This process consists of six phases: Interphase, prophase, prometaphase, metaphase, anaphase, and telophase. Each phase in the process was described on a separate slide, accompanied by a drawing depicting that phase. That the drawings were relevant for learning had been established in prior research with these learning materials, which showed that the expository text accompanied by pictures led to better learning outcomes than the text alone (Scheiter, Schüler, Gerjets, Huk, & Hesse, 2014; Schüler, Scheiter, & Gerjets, 2013).

The experiment comprised three conditions: In the control condition only the relevant text and the picture were shown on each slide (see Figure 1). In the separated unnecessary text condition, the pictures were additionally accompanied by a description of the visuo-spatial appearance of the components shown in the picture (see Figure 1). In the integrated condition, the same information as in the separated condition was presented, but the unnecessary text describing the pictures was integrated with the pictures (see Figure 2).

On average, the relevant text consisted of 77 words on each slide (range 72-82), while the unnecessary text consisted of 39 words per slide (range 28-46). The location of the picture and the unnecessary text was varied (left- or right-hand side of the screen) between subjects to control for potential bias in attention as a result of reading direction. The learning materials were system-paced and the available time was the same in each condition, so participants could not compensate for time spent on processing the unnecessary text by investing more time on the materials overall. A small user-paced pilot (n = 8) was used to determine the presentation time per slide. We calculated the average time those eight participants spent on each slide, and used this as the presentation time of the corresponding slide in the current experiment (slide 1: 65s, slide 2: 120s, slide 3: 110s, slide 4: 97s, slide 5: 107s, slide 6: 77s). It was not possible for participants to go back to a previously presented slide.

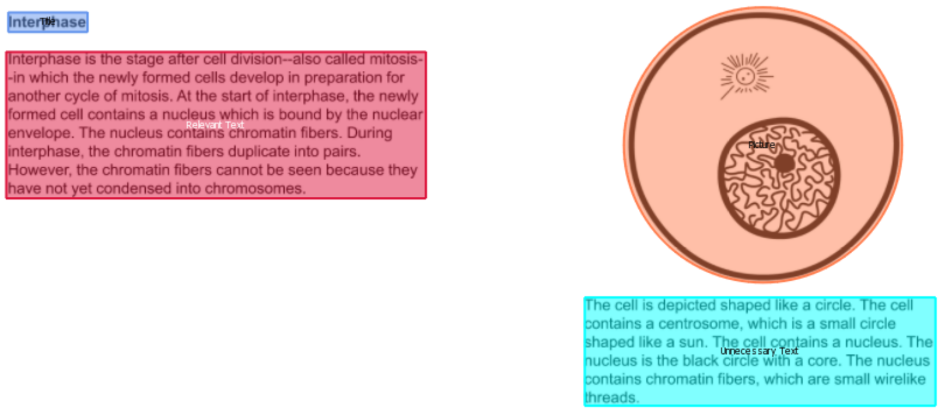


Figure 1. Example slide for the separated and control conditions with the Areas of Interest used in Experiment 2. The description of the components in the picture (unnecessary information) is presented underneath the picture. In the control condition, this description is not present.

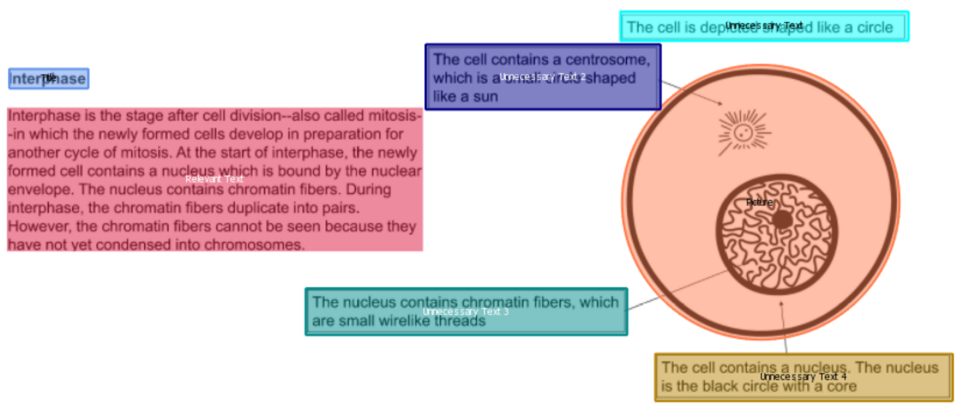


Figure 2. Example slide for the integrated condition with the Areas of Interest used in Experiment 2. The description of the components in the picture is integrated in the picture.

Cloze test. Knowledge about the studied mitosis phase was tested immediately after each slide, using a cloze test in which participants were presented with four short sentences from the relevant text they had just studied, with one or two keyword(s) omitted (e.g., The nucleus of the newly formed cell is bound by the __). Participants were asked to fill in the blanks by typing the correct answer into the answer box.

Invested mental effort. Participants were asked to indicate how much effort they invested in learning the content of each preceding slide on a nine-point rating scale (Paas, 1992), ranging from one (*extremely low effort*) to nine (*extremely high effort*). Moreover, participants were asked to indicate how much effort they invested in answering the cloze test after each slide, using the same nine-point scale.

Picture test. Because processing the unnecessary text might have gone at the expense of processing the pictures, we also tested participants' knowledge of the pictures at the end of the experiment. To do so, we used a multiple-choice test consisting of seven items. In six items (present in a random order), participants were presented with a picture of one of the phases, and had to choose which phase it depicted, from six alternatives. In the seventh item, participants saw all six phases depicted on the screen, and had to indicate the correct order of the pictures (i.e., according to the phases of mitosis) from six possible answers.

Procedure

Participants were either tested individually or with two participants simultaneously. First, the prior knowledge test was administered and participants were asked to fill in their age and gender. After this test, participants learned about mitosis with a slideshow consisting of six slides, each depicting a different phase in the process, under one of the three conditions (see Figures 1 and 2). After each slide, participants first had to indicate how much effort they invested in studying the preceding slide, then fill in the cloze test, and then indicate how much effort they invested in answering the cloze test questions. After the learning phase, participants had to fill in the picture test. In total, the experiment took approximately 20 to 30 minutes, and it was administered without breaks.

Scoring

For all multiple-choice questions, participants received one point when they gave the correct answer and no points when they gave the wrong answer. Thus, participants could score a maximum of four points on the pretest, and seven points on the picture test. For the cloze test questions, participants were awarded 1 point if the correct answer was given, 0.5 points when the answer was partially correct, and 0 points when they did not provide an answer or if it was completely wrong. Thus, participants could score zero to four points per cloze test after each slide. A random subset of the cloze test data (10.4%) was scored by a second rater, and interrater reliability was high ($\kappa = .91$).

Results

As mentioned in the materials section, we controlled for reading direction by counterbalancing the location of the picture and the unnecessary text (hereafter called PUT). As a check revealed that PUT-location seemed to influence the dependent variables, we included it as a factor in the analyses. In the analyses of the effects of task experience (i.e., on cloze test performance and invested mental effort during learning and in the cloze test), we made the distinction between lower (slide 1 to 3) and higher task experience (slide 4 to 6). When the sphericity assumption was violated, we report the results after Greenhouse-Geisser correction. We used partial eta-squared and Cohen's *d* as measures of effect size; both can be interpreted in terms of small ($\eta_p^2 \sim .01$, $d \sim 0.2$), medium ($\eta_p^2 \sim .06$, $d \sim 0.5$), and large ($\eta_p^2 \sim .14$, $d \sim 0.8$) effect sizes (Cohen, 1988). Moreover, when post-hoc follow-up tests were performed, we used a Bonferroni correction (i.e., multiplying the *p*-value with the number of tests performed).

Prior knowledge

Performance on the prior knowledge test is presented in Table 1, and was analyzed with a 3 x 2 ANOVA with condition (separated, integrated, or control) and PUT location (left or right) as between-subjects factors. The analyses revealed no effect of condition, $F < 1$, no effect of PUT location, $F(1, 88) = 3.00$, $p = .087$, $\eta_p^2 = .03$, and no interaction, $F(2, 88) = 1.81$, $p = .169$, $\eta_p^2 = .04$. Hence, there were no significant differences in prior knowledge among conditions.

Table 1. *Mean (and SD) Performance on the Prior Knowledge Test (max. = 4) and Picture Test (max. = 7) as a Function of Condition and PUT Location in Experiment 1 and 2.*

		Prior knowledge test		Picture test	
		Exp. 1	Exp. 2	Exp. 1	Exp. 2
PUT left	Control	2.76 (1.09)	2.26 (1.10)	4.59 (1.77)	4.89 (1.70)
	Integrated	2.19 (0.98)	2.50 (1.05)	4.69 (1.82)	4.95 (1.79)
	Separated	2.80 (1.08)	2.23 (1.23)	4.67 (1.99)	5.32 (1.43)
	Total	**2.58 (1.07)**	**2.33 (1.12)**	**4.65 (1.82)**	**5.07 (1.62)**
PUT right	Control	2.13 (1.13)	2.37 (0.96)	4.07 (1.22)	5.79 (1.44)
	Integrated	2.40 (1.06)	2.55 (1.19)	5.20 (2.01)	5.40 (1.64)
	Separated	2.06 (1.12)	2.41 (1.14)	4.81 (1.56)	5.23 (1.66)
	Total	**2.20 (1.09)**	**2.44 (1.09)**	**4.70 (1.66)**	**5.46 (1.58)**
Total	Control	2.47 (1.14)	2.32 (1.02)	4.34 (1.54)	5.34 (1.62)
	Integrated	2.29 (1.01)	2.53 (1.11)	4.94 (1.90)	5.18 (1.71)
	Separated	2.42 (1.15)	2.32 (1.18)	4.74 (1.75)	5.27 (1.53)
	Total	**2.39 (1.09)**	**2.39 (1.10)**	**4.67 (1.73)**	**5.26 (1.60)**

Cloze test

The average performance on the first three (lower experience) and last three cloze tests (higher experience) is presented in Table 2. We performed a 3 x 2 x 2 mixed ANOVA with between-subjects factors condition (separated, integrated, or control) and PUT location (left or right) and within-subjects factor task experience (lower or higher) on these data. The analysis revealed a main effect of condition, $F(2, 88) = 5.06$, $p = .008$, $\eta_p^2 = .10$, with follow-up tests showing that performance in the separated condition ($M = 1.19$, $SD = 0.62$) was significantly lower than performance in the control condition ($M = 1.71$, $SD = 0.65$), $p = .005$, $d = 0.82$. Performance in the integrated condition ($M = 1.45$, $SD = 0.64$) did not significantly differ from performance in the control condition, $p = .349$, $d = 0.39$, or the separated condition $p = .319$, $d = 0.42$. Moreover, the analysis revealed a main effect of task experience, $F(1, 88) = 10.20$, $p = .002$, $\eta_p^2 = .10$, indicating that cloze test performance was higher on the first three slides, when participants had lower task experience ($M = 1.55$, $SD = 0.74$), than on the last three slides, when they had more task experience ($M = 1.36$, $SD = 0.74$). We found no main effect of PUT location, $F < 1$, but the analysis did reveal an interaction between task experience and PUT location, $F(1, 88) = 11.11$, $p =$

.001, $\eta_p^2 = .11$. Two-tailed independent samples t-tests showed that the location of the unnecessary information (left, *M* = 1.51, *SD* = 0.73; right, *M* = 1.59, *SD* = 0.75) did not influence cloze test performance when participants had lower task experience, $t(92) = 0.53$, $p > .999$, $d = 0.11$, while participants with more task experience seemed to perform better when the unnecessary text was presented left (*M* = 1.52, *SD* = 0.78) than when it was presented right (*M* = 1.20, *SD* = 0.66), $t(92) = 2.18$, $p = .064$, $d = 0.45$. We found no interaction between PUT location and condition, $F < 1$. More relevant for our hypotheses, we found no interaction between task experience and condition, $F(2, 88) = 1.65$, $p = .198$, $\eta_p^2 = .04$. However, we did find a three way interaction between condition, task experience, and PUT location, $F(2, 88) = 3.51$, $p = .034$, $\eta_p^2 = .07$. This three way interaction presumably arose because initial performance differences between the control and separated condition diminished when participants gained task experience, but only when the unnecessary information was presented on the right-hand side of the screen (see Figure 3). We did no predict this three-way interaction based on our theoretical framework, and the pattern of results was not in line with it. This is because performance in the control condition diminished while reasoning from our theoretical framework one would predict the performance level to remain constant in the control condition, whereas it ought to increase in the separated condition.

Table 2. *Mean (and SD) Cloze Test Performance (max. = 4) as a Function of Condition, PUT location, and Task Experience in Experiment 1 and 2.*

		Experiment 1		Experiment 2	
		Task Experience		Task Experience	
		Low	High	Low	High
PUT left	Control	1.72 (0.71)	1.76 (0.69)	1.48 (0.64)	1.31 (0.70)
	Integrated	1.48 (0.71)	1.67 (0.77)	1.43 (0.65)	1.38 (0.71)
	Separated	1.30 (0.74)	1.09 (0.73)	1.50 (0.50)	1.39 (0.59)
	Total	**1.51 (0.73)**	**1.52 (0.78)**	**1.47 (0.59)**	**1.36 (0.66)**
PUT right	Control	2.02 (0.72)	1.32 (0.75)	1.58 (0.83)	1.34 (1.06)
	Integrated	1.48 (0.72)	1.18 (0.61)	1.29 (0.73)	1.32 (0.61)
	Separated	1.28 (0.66)	1.09 (0.64)	1.56 (0.77)	1.60 (0.87)
	Total	**1.59 (0.75)**	**1.20 (0.66)**	**1.48 (0.77)**	**1.43 (0.86)**
Total	Control	1.86 (0.72)	1.56 (0.74)	1.53 (0.73)	1.32 (0.89)
	Integrated	1.48 (0.70)	1.43 (0.73)	1.36 (0.69)	1.35 (0.66)
	Separated	1.29 (0.69)	1.09 (0.67)	1.53 (0.64)	1.50 (0.74)
	Total	**1.55 (0.74)**	**1.36 (0.74)**	**1.47 (0.69)**	**1.39 (0.76)**

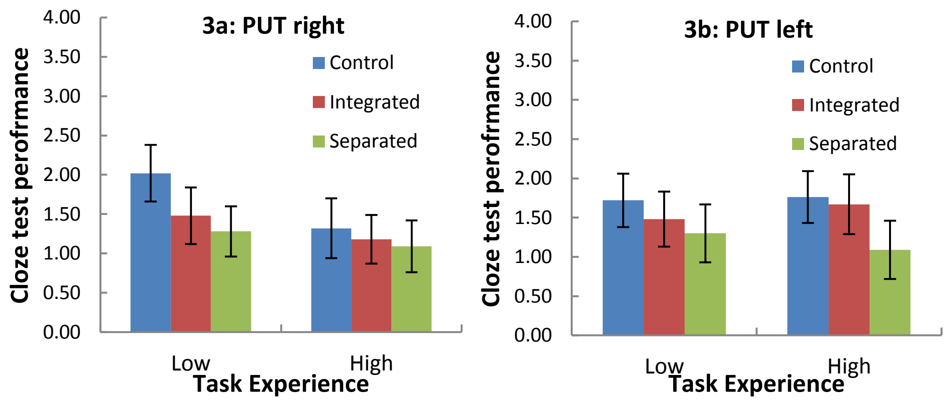


Figure 3. Mean Cloze Test Performance in Experiment 1 (max. = 4) as a Function of Condition and Task Experience, when the PUT location was Right (Figure 3a) or Left (figure 3b).

Mental effort

Self-reported invested mental effort during the learning phase and the cloze test are presented in Tables 3 and 4. We performed 3 x 2 x 2 mixed ANOVAs with between-subjects factors condition (separated, integrated, or control) and PUT location (left or right) and within-subjects factor task experience (lower or higher) on the invested mental effort data obtained during the learning phase and the cloze test. For the invested mental effort during the learning phase, this analysis revealed no main effect of condition, $F < 1$, and no main effect of PUT location, $F(1, 88) = 2.41$, $p = .124$, $\eta_p^2 = .03$. The analysis did reveal a main effect of task experience, $F(1, 88) = 6.16$, $p = .015$, $\eta_p^2 = .07$, indicating that participants invested more mental effort when they had lower task experience, ($M = 6.73$, $SD = 1.09$) than when they had more task experience, ($M = 6.52$, $SD = 1.21$). There were no significant interactions, all Fs < 1.

Table 3. *Mean (and SD) Invested Mental Effort (max. = 9) during the Learning Phase as a Function of Condition, PUT location, and Task Experience in Experiment 1 and Experiment 2.*

		Experiment 1		Experiment 2	
		Task Experience		Task Experience	
		Low	High	Low	High
PUT left	Control	7.00 (0.89)	6.86 (1.17)	6.58 (1.05)	6.61 (1.36)
	Integrated	6.88 (1.11)	6.80 (1.11)	6.40 (1.16)	6.10 (1.24)
	Separated	6.78 (0.72)	6.47 (0.88)	6.15 (1.05)	6.09 (1.03)
	Total	**6.89 (0.91)**	**6.72 (1.06)**	**6.37 (1.08)**	**6.26 (1.22)**
PUT right	Control	6.62 (0.77)	6.47 (1.27)	6.60 (1.09)	6.19 (1.21)
	Integrated	6.44 (1.52)	6.07 (1.62)	6.50 (1.24)	6.30 (1.39)
	Separated	6.65 (1.36)	6.44 (1.17)	6.58 (0.89)	6.45 (0.85)
	Total	**6.57 (1.24)**	**6.33 (1.34)**	**6.56 (1.06)**	**6.32 (1.15)**
Total	Control	6.82 (0.85)	6.68 (1.21)	6.59 (1.05)	6.40 (1.29)
	Integrated	6.67 (1.32)	6.44 (1.40)	6.45 (1.18)	6.20 (1.30)
	Separated	6.71 (1.08)	6.45 (1.02)	6.36 (0.99)	6.27 (0.95)
	Total	**6.73 (1.09)**	**6.52 (1.21**	**6.46 (1.07)**	**6.29 (1.18)**

For the invested mental effort during the cloze test, this analysis revealed no main effect of condition, $F < 1$, no main effect of PUT location, $F(2, 88) = 1.82$, $p = .181$, $\eta_p^2 = .02$, and no main effect of task experience, $F < 1$. The analysis revealed no interaction between PUT location and condition, $F(2, 88) = 2.29$, $p = .107$, $\eta_p^2 = .05$, PUT location and task experience, $F(2, 88) = 1.93$, $p = .168$, $\eta_p^2 = .02$, task experience and condition, $F < 1$, nor a three-way interaction condition x PUT location x task experience, $F(2, 88) = 1.15$, $p = .323$, $\eta_p^2 = .03$.

Table 4. *Mean (and SD) Invested Mental Effort (max. = 9) during the Cloze Test as a Function of Condition, PUT location, and Task Experience in Experiment 1 and Experiment 2.*

		Experiment 1		Experiment 2	
		Task Experience		Task Experience	
		Low	High	Low	High
PUT left	Control	6.88 (1.25)	6.86 (1.17)	6.26 (1.39)	6.37 (1.46)
	Integrated	6.90 (0.84)	6.89 (0.84)	5.77 (1.21)	5.62 (1.18)
	Separated	6.02 (1.07)	6.02 (1.06)	5.91 (1.36)	5.82 (1.07)
	Total	**6.62 (1.12)**	**6.72 (1.06)**	**5.97 (1.36)**	**5.92 (1.26)**
PUT right	Control	6.04 (1.01)	6.47 (1.27)	5.98 (1.09)	5.74 (1.11)
	Integrated	6.09 (1.46)	6.09 (1.46)	6.02 (1.39)	6.08 (1.45)
	Separated	6.40 (1.30)	6.44 (1.27)	6.05 (1.05)	5.98 (0.98)
	Total	**6.18 (1.25)**	**6.32 (1.34)**	**6.02 (1.16)**	**5.94 (1.18)**
Total	Control	6.49 (1.21)	6.68 (1.21)	6.12 (1.24)	6.05 (1.32)
	Integrated	6.51 (1.23)	6.44 (1.40)	5.89 (1.29)	5.85 (1.32)
	Separated	6.22 (1.19)	6.45 (1.02)	5.98 (1.26)	5.90 (1.02)
	Total	**6.40 (1.20)**	**6.52 (1.21)**	**5.99 (1.26)**	**5.93 (1.21)**

Picture test

Performance on the picture test is shown in see Table 1, and was analyzed with a 3 x 2 ANOVA with condition (separated, integrated, or control) and PUT location (left or right) as between-subjects factors. The analysis revealed no effect of condition, $F(2, 88) = 1.02$, $p = .366$, $\eta_p^2 = .02$, no effect of PUT location, $F < 1$, nor an interaction effect, $F < 1$.

Discussion

Surprisingly, we did not find a consistent negative effect of unnecessary information on (initial) learning. Unnecessary information only had a negative effect on cloze test performance when it was presented separated from the relevant information, not when it was integrated. This finding was qualified, however, by a three-way interaction between condition, task experience, and PUT location, suggesting that initial performance differences between the control and separated condition diminished when participants gained task experience, but only when the unnecessary information was presented on the right-hand side of the screen. In contrast to our expectations, however, the reduced difference between the conditions seemed to result from a decline in performance in the control condition rather than an increase in performance in the separated condition. In sum, our hypotheses regarding task experience were not confirmed in Experiment 1. To get more insight into how students process the unnecessary text and whether this changes over time (i.e., with increasing task experience), Experiment 2 replicated Experiment 1, using eye-tracking methodology.

Experiment 2

In Experiment 2, we employed eye tracking to investigate how much attention participants devoted to the unnecessary information in the separated and integrated conditions, and whether they would start to ignore it over time. We hypothesized that the unnecessary text would initially attract attention, but with increasing task experience, participants would start to ignore the unnecessary text and allocate less attention to it. This should result in 1) shorter fixation duration on the unnecessary text, and longer fixation duration on the relevant text and picture with increasing task experience; and 2) more transitions between relevant information sources, and less transitions between relevant and unnecessary information sources with increasing task experience.

Method

Participants and design

Participants were 133 German University students (M_{age} = 21.28 years, SD = 2.34 years; 107 female) who participated for course credit or a small fee of 5 euro. All participants had normal or corrected-to-normal vision. One participant indicated after completing the experiment that he/she wanted to retract his/her data. For six participants the data on how long they spend on each slide indicated that they had skipped parts of the learning phase. Furthermore, due to a randomization error, four participants participated in two conditions of the Experiment (i.e., they saw each slide twice, in two different conditions). The data of these eleven participants were excluded from all future analyses, resulting in a sample of 122 participants (M_{age} = 21.06 years, *SD* = 2.28 years; 98 female). Participants were randomly assigned to one of the three conditions: control (*n* = 38), separated (*n* = 44), and integrated (*n* = 40). Again, within conditions PUT location was varied: for half of the participants the picture and unnecessary text was presented at the right, while for the other half the picture and unnecessary text was presented at the left.

Apparatus and materials

The materials were identical to those of Experiment 1. The materials were presented in SMI Experiment Center (Version 3.6; SensoMotoric Instruments), on a monitor with a resolution of 1920 x 1080 pixels. Participants' eye movements were recorded using SMI RED 250 Mobile eye trackers (SensoMotoric Instruments) that record binocularly at 250 Hz using SMI iView software (Version 2.8; SensoMotoric Instruments). The data was subsequently analyzed using BeGaze software (Version 3.7; SensoMotoric Instruments).

Procedure

The procedure was similar to that of Experiment 1, only in Experiment 2 participants were tested individually, or in groups of up to three participants simultaneously, and their eye movements were recorded during the learning phase. At the start of the Experiment, participants were seated in front of a mobile eye tracker, with their head approximately 60 cm from the monitor. After a short introduction and the prior knowledge test, the eye tracker was calibrated using a thirteen-point calibration plus four-point validation procedure, and participants were instructed to move as little as possible. The Experiment lasted around twenty minutes and was administered without breaks.

Data analysis

For the eye tracking analyses, we first checked the accuracy of calibration, which was sufficient for all participants (i.e., no deviations from the four validation points of more than 1° visual angle). We then checked the tracking ratio (i.e., the percentage of time for which the eye tracker actually measured the eye movements) for each participant. We had to exclude 22 participants (control: *n* = 8; separated: *n* = 5; integrated *n* = 9) because their tracking ratio was below 70%. The final sample (*n* = 100) had an average tracking ratio of 92.03% (*SD* = 6.75%), with a mean calibration accuracy of 0.28° (*SD* = 0.15°) and was distributed across the conditions as follows: control (n = 30), separated (n = 39), and integrated (n = 31).

For the eye tracking analyses we defined fixations using a 40 °/s velocity threshold and a minimal duration of 100 ms (cf. Holmqvist et al., 2011). On each slide in each condition, we created areas of interest (AoIs) for the picture, for the relevant text and for the title (see Figure 1). In the separated condition, we defined one extra AoI for the unnecessary text (see Figure 1), while in the integrated condition we created additional AoIs for each text block (see Figure 2). In this condition, in the first three phases, there were four unnecessary text blocks, while the slides in the last three phases had three unnecessary text blocks.

As a measure of attention to the different AoIs, we used fixation time. Because presentation time between slides was different (see the Learning materials section of Experiment 1), and the size of the AoIs was different between conditions (see Figure 1 and 2), we had to calculate a relative measure of fixation time. We did so by dividing the fixation time on each AoI by the percentage of the screen covered by that AoI to control for the size of the AoI. We then divided this value by the total fixation time on that slide in seconds (i.e., the sum of all fixations on the different AoIs and white space), to control for the differences in presentation duration and tracking ratio.
To measure integration of the different sources of information (i.e., relevant text, picture, unnecessary text), we used transitions between the different AoIs. We defined three types of transitions: relevant-picture transitions, which are transitions between the picture and relevant text and vice versa; unnecessary-relevant transitions, which are transitions between the unnecessary and relevant text and vice versa; and unnecessary-picture transitions, which are transitions between the unnecessary text and the picture and vice versa. The unnecessary-relevant and unnecessary-picture transitions can only be calculated for the two unnecessary text conditions. To control for differences in presentation duration between slides, we divided the number of transitions by the total fixation time on that slide in seconds (i.e., the same value we used in the fixation time measure).

Results

The data on prior knowledge, cloze test performance, invested mental effort, and picture test performance are analyzed with the same ANOVAs as in Experiment 1.

Prior knowledge

Performance on the prior knowledge test is presented in Table 1. The analysis revealed no effect of condition, $F < 1$, no effect of PUT location, $F < 1$, nor an interaction, $F < 1$.

Cloze test

On the cloze test performance (see Table 3), the analyses revealed no main effect of condition, $F < 1$, no main effect of PUT location, $F < 1$, and no main effect of task experience, $F(2, 116) = 2.50$, $p = .116$, $\eta_p^2 = .02$. Furthermore, we found no interaction between condition and task experience, $F(2, 116) = 1.35$, $p = .262$, $\eta_p^2 = .02$, between task experience and PUT location, $F < 1$, or between condition and PUT location, $F < 1$. Finally, we did not find any evidence for the three-way interaction observed in Experiment 1, $F < 1$.

Invested mental effort

The invested mental effort during learning is presented in Table 3. The analysis revealed no main effect of condition, $F < 1$, and no main effect of PUT location, $F < 1$. However, we did find a main effect of task experience, $F(2, 116) = 4.45$, $p = .037$, $\eta_p^2 = .04$, indicating that invested mental effort during learning was higher on the first three slides, when participants had lower task experience ($M = 6.46$, $SD = 1.07$) than on the last three slides when they had more task experience ($M = 6.29$, $SD = 1.18$). We found no interactions, all Fs < 1.

The analysis on the invested mental effort during the cloze test (see Table 4) revealed no main effect of condition, $F < 1$, no main effect of PUT location, $F < 1$, and no main effect of task experience, $F < 1$. Furthermore, we found no interactions between condition and task experience, $F < 1$, between PUT location and task experience, $F < 1$, between condition and PUT location, $F(2, 116) = 1.27$, $p = .258$, $\eta_p^2 = .02$, nor a three-way interaction, $F(2, 116) = 1.16$, $p = .317$, $\eta_p^2 = .02$.

Picture test

Regarding the picture test (see Table 1), the analysis showed no effect of condition, $F < 1$, no effect of PUT location, $F(2, 116) = 2.04$, $p = .156$, $\eta_p^2 = .02$, or an interaction, $F < 1$.

Eye movement data

The eye movement data were analyzed in two steps. First we tested whether the presence and layout of unnecessary information leads to differences in attention towards unnecessary and relevant information. Secondly, we tested whether the presence and layout of unnecessary information would lead to differences in integration of text and pictures. To do so, we performed a 3 x 2 x 2 mixed ANOVA with between-subjects factors condition (separated, integrated, or control) and PUT location (left or right) and within-subjects factor task experience (lower or higher) on the fixation time on the relevant text and the picture. On the fixation time on the unnecessary text, we performed a 2 x 2 x 2 mixed ANOVA with between-subjects factors condition (separated or integrated) and PUT location (left or right) and within-subjects factor task experience (lower or higher). The data on the relevant-picture, unnecessary-picture, and unnecessary-relevant transitions are analyzed with non-parametric tests as the assumptions of normality were violated.

Fixation time. The data on the fixation time (corrected for AoI size and total fixation time, see data analysis section) on the *unnecessary text* are presented in Table 5. The analysis revealed a significant main effect of condition, $F(1, 66) = 21.14$, $p < .001$, $\eta_p^2 = .24$, indicating that participants in the integrated condition ($M = 17.57$, $SD = 7.80$) spent less time fixating on the unnecessary text than participants in the separated condition ($M = 28.62$, $SD = 11.53$). Moreover, the analysis revealed a significant effect of task experience, $F(1, 66) = 6.98$, $p = .010$, $\eta_p^2 = .10$, indicating that participants spent less time fixating on the unnecessary text on the last three slides, after they had gained task experience ($M = 21.84$, $SD = 14.59$), compared to on the first three slides, when they had lower task experience ($M = 25.60$, $SD = 11.30$). The analysis showed no effect of PUT location, $F(1, 66) = 3.58$, $p = .063$, $\eta_p^2 = .05$, and no interaction effects, smallest $p = .056$, $\eta_p^2 = .05$.

For the fixation time on the *relevant text* (see Table 5), our analysis revealed a significant main effect of condition, $F(2, 94) = 31.93$, $p < .001$, $\eta_p^2 = .41$. Follow-up tests showed that participants in the integrated condition ($M = 49.31$, $SD = 7.56$; $p < .001$, $d = 1.81$) and separated condition ($M = 50.20$, $SD = 7.51$; $p < .001$, $d = 1.68$), fixated less on the relevant text than participants in the control condition ($M = 61.90$, $SD = 6.33$). Time spent fixating on the relevant text did not differ significantly between the integrated and separated conditions, $p > .999$, $d = 0.12$. We found a significant effect of task experience, $F(2, 94) = 40.56$, $p < .001$, $\eta_p^2 = .30$, indicating that participants spent more time fixating the relevant text after they gained task experience ($M =$

56.40, *SD* = 10.26), compared to when they had lower task experience (*M* = 50.46, *SD* = 10.11). Furthermore, the analysis revealed a main effect of PUT location, $F(2, 94) = 17.47$, $p < .001$, $\eta_p^2 = .16$, indicating that participants attended more to the relevant text when the PUT was on the left (*M* = 56.64, *SD* = 8.44) than when it was on the right side of the screen (*M* = 50.48, *SD* = 8.64). There were no interaction effects, smallest $p = .070$, $\eta_p^2 = .06$.

Regarding the fixation time on the *picture* (see Table 5), we found no significant effect of condition, $F(2, 94) = 1.69$, $p = .189$, $\eta_p^2 = .04$, or task experience, $F(2, 94) = 1.67$, $p = .200$, $\eta_p^2 = .02$. However, the analysis revealed a main effect of PUT location, $F(2, 94) = 8.11$, $p = .005$, $\eta_p^2 = .08$, indicating that participants fixated more on the picture when the PUT was presented on the right (*M* = 6.77, *SD* = 3.12) than when it was presented on the left side of the screen (*M* = 5.14, *SD* = 2.98). The analyses revealed no interactions, smallest $p = .116$, $\eta_p^2 = .05$.

Table 5. *Mean (and SD) Fixation Duration on the Different AoIs as a Function of Condition, PUT location, and Task Experience.*

		Unnecessary Text		Relevant text		Picture	
		Task Experience		Task Experience		Task Experience	
		Low	High	Low	High	Low	High
PUT left	Control			63.63 (6.90)	66.02 (6.02)	5.09 (3.45)	5.91 (4.86)
	Integrated	17.49 (4.79)	13.63 (9.13)	49.44 (5.99)	57.64 (8.99)	5.43 (2.81)	4.87 (2.99)
	Separated	30.60 (11.87)	20.82 (14.18)	46.41 (7.54)	57.12 (9.27)	4.69 (2.67)	4.93 (3.23)
	Total	**24.86 (11.43)**	**17.68 (12.58)**	**53.03 (10.23)**	**60.24 (9.06)**	**5.04 (2.94)**	**5.24 (3.74)**
PUT right	Control			56.92 (7.28)	60.17 (7.37)	7.74 (3.26)	7.47 (4.10)
	Integrated	20.52 (10.65)	17.92 (8.54)	42.83 (8.40)	48.80 (9.04)	6.41 (3.09)	8.33 (3.98)
	Separated	30.85 (9.81)	31.37 (17.16)	46.46 (7.61)	51.27 (10.27)	5.50 (3.63)	5.98 (3.11)
	Total	**26.23 (11.32)**	**25.36 (15.38)**	**48.09 (9.49)**	**52.86 (10.10)**	**6.40 (3.42)**	**7.15 (3.75)**
Total	Control			60.50 (7.74)	63.29 (7.20)	6.33 (3.57)	6.63 (4.51)
	Integrated	19.15 (8.53)	15.99 (8.93)	45.82 (8.02)	52.79 (9.93)	5.97 (2.96)	6.77 (3.92)
	Separated	30.73 (10.66)	26.50 (16.53)	46.44 (7.48)	53.97 (10.13)	5.12 (3.21)	5.50 (3.17)
	Total	**25.60 (11.30)**	**21.84 (14.59)**	**50.46 (10.11)**	**56.40 (10.26)**	**5.75 (3.26)**	**6.23 (3.85)**

Transitions. On the *unnecessary-picture transitions* (i.e., transitions between the unnecessary text and the picture; see Table 6), a Mann-Whitney test revealed a main effect of condition, $U = 314.00$, $p = .001$, indicating that participants in the integrated condition (*Mdn* = 0.11, *Range* = 0.34) made more unnecessary-picture transitions than participants in the separated condition (*Mdn* = 0.06, *Range* = 0.34). A Wilcoxon Signed Ranks test showed a significant effect of task experience, $Z = 4.75$, $p < .001$, indicating that participants made fewer unnecessary-picture transitions after they gained task experience (*Mdn* = 0.07, *SD* = 0.49) than when they had little task experience (*Mdn* = 0.11, *Range* = 0.52). Finally, a Mann-Whitney test revealed no effect of PUT location, $U = 558.00$, $p = .556$.

Regarding the *unnecessary-relevant transitions* (i.e., transitions between the unnecessary text and the relevant text; see Table 6), the analysis revealed a main effect of condition, $U = 353.00$, $p = .003$, indicating that participants in the integrated condition (*Mdn* = 0.021, *Range* = 0.07) made more unnecessary-relevant transitions than participants in the separated condition (*Mdn* = 0.013, *Range* = 0.13). We again found a main effect of task experience, $Z = 3.64$, $p < .001$, indicating that participants made fewer unnecessary-relevant transitions after they gained task experience (*Mdn* = 0.013, *Range* = 0.01) than when they had lower task experience (*Mdn* = 0.018, *Range* = 0.02). Furthermore, the analysis revealed a main effect of PUT location, $U = 301.00$, $p < .001$, showing that participants made more unnecessary-relevant transitions when the PUT was presented at the right (*Mdn* = 0.023, *Range* = 0.13) than when it was presented at the left (*Mdn* = 0.011, *Range* = 0.04).

Regarding the *relevant-picture transitions* (i.e., transitions between the relevant text and the picture; see Table 6), a Kruskall-Wallis test revealed a significant effect of condition, χ^2 (2) = 11.56, $p = .003$. Follow-up Mann-Whitney tests showed that both participants in the integrated (*Mdn* = 0.05, *Range* = 0.45; $U = 247.00$, $p = .006$) and separated condition (*Mdn* = 0.07, *Range* = 0.39; $U = 359.00$, $p = .018$) made significantly fewer relevant-picture transitions than participants in the control condition (*Mdn* = 0.12, *Range* = 0.40). Participants in the integrated and separated conditions did not differ in the number of relevant-picture transitions, $U = 549.00$, $p > .999$. We found a significant effect of task experience, $Z = 3.36$, $p = .001$, indicating that participants made more relevant-picture transitions after they gained task experience (*Mdn* = 0.09, *Range* = 0.79), compared to when they had little task experience (*Mdn* = 0.07, *Range* = 0.51). Finally, the analysis revealed an effect of PUT location, $U = 904.00$, $p = .018$, showing that participants made more relevant-picture transitions when the PUT was presented at the right (*Mdn* = 0.09, *Range* = 0.45) then when it was presented at the left (*Mdn* = 0.05, *Range* = 0.37).

Table 6. *Median (and Range) number of Transitions between the Different AoIs as a Function of Condition, PUT location, and Task Experience.*

		Relevant-Picture		Unnecessary-Picture		Unnecessary-Relevant	
		Task Experience		Task Experience		Task Experience	
		Low	High	Low	High	Low	High
PUT left	Control	0.12 (0.44)	0.11 (0.45)				
	Integrated	0.03 (0.09)	0.04 (0.25)	0.11 (0.19)	0.07 (0.16)	0.018 (0.03)	0.012 (0.05)
	Separated	0.03 (0.35)	0.05 (0.25)	0.11 (0.52)	0.05 (0.13)	0.008 (0.05)	0.000 (0.03)
	Total	**0.04 (0.47)**	**0.05 (0.46)**	**0.11 (0.52)**	**0.06 (0.18)**	**0.012 (0.05)**	**0.009 (0.05)**
PUT right	Control	0.13 (0.29)	0.11 (0.69)				
	Integrated	0.08 (0.49)	0.11 (0.78)	0.13 (0.20)	0.10 (0.48)	0.034 (0.11)	0.018 (0.08)
	Separated	0.08 (0.38)	0.07 (0.40)	0.07 (0.42)	0.07 (0.26)	0.016 (0.16)	0.016 (0.10)
	Total	**0.08 (0.51)**	**0.10 (0.78)**	**0.10 (0.42)**	**0.08 (0.49)**	**0.025 (0.16)**	**0.017 (0.01)**
Total	Control	0.12 (0.44)	0.11 (0.75)				
	Integrated	0.05 (0.51)	0.06 (0.79)	0.12 (0.20)	0.08 (0.48)	0.024 (0.11)	0.016 (0.08)
	Separated	0.06 (0.38)	0.06 (0.42)	0.08 (0.52)	0.07 (0.26)	0.010 (0.16)	0.009 (0.10)
	Total	**0.07 (0.51)**	**0.09 (0.79)**	**0.11 (0.52)**	**0.07 (0.49)**	**0.018 (0.16)**	**0.013 (0.10)**

Discussion

The results of Experiment 2 are mixed. While the eye-tracking measures supported our hypotheses, indicating that attention towards the unnecessary text waned with increasing task experience, this did not affect cloze test performance. In contrast to our hypothesis, presentation of unnecessary text did not initially hamper learning about the process of mitoses, regardless of whether the unnecessary text was presented integrated in, or separated from the picture. Because we found no initial negative effect of the unnecessary text on learning, the question of whether task experience would reduce or eliminate that negative effect could not be answered.

Next to the main finding of diminishing attention to the unnecessary text in favour of attention to the essential textual information, the eye-tracking analyses suggested an effect of screen location of the unnecessary text and picture. When the relevant text was on the right-hand side of the screen (and the picture + unnecessary text on the left), participants attended more to it than when it was on the left-hand side of the screen. A possible explanation is that this might be due to the fact that participants looked at the *center* of the screen at the beginning of each new slide because the mental effort question on the preceding slide was presented in the center. When attention is centrally located, one may be inclined (because of Western reading direction) to process information on the right-hand side of the screen first.

General Discussion

According to well-known principles in multimedia learning, the presentation of extraneous information should be avoided, because it hinders learning (i.e., the coherence principle, Mayer & Fiorella, 2014; the redundancy principle, Kalyuga & Sweller, 2014). However, recent research with irrelevant pictorial information demonstrated that students may learn to ignore extraneous information with task experience, at which point it no longer negatively affects their learning (Rop, Van Wermeskerken et al., in press). The present study aimed to examine whether these findings would extend to extraneous information that is textual, and unnecessary rather than irrelevant. Moreover, we investigated the role of the layout of the unnecessary textual information: we expected that it would be harder for students to (learn to) ignore unnecessary text when it is presented spatially integrated in a relevant picture, as compared to spatially separated from the picture.

The eye-movement data collected in Experiment 2 showed that the unnecessary textual information was processed by students, and more interestingly, that they seemed to start ignoring the unnecessary information with increasing task experience. That is, participants paid less attention to the unnecessary text and made less transitions between the unnecessary and essential information on the later slides, after they had gained some experience with the task. These results are in line with the findings by Rop, Van Wermeskerken et al. (in press), who showed that learners start to ignore pictorial, obviously irrelevant, and separated extraneous information. The present study shows that these results also apply when the extraneous information is textual, unnecessary rather than irrelevant, and when it is integrated or separated with relevant information. This provides further evidence that learners adapt their study strategy and start to focus less on extraneous and more on essential information once they gain experience with a task, which is relevant information for instructional designers.

Although the results also implied that more attention was paid to the essential text on the later slides, this change in study strategy did not lead to improvements in test performance (improvements which were observed by Rop, Van Wermeskerken et al., in press). Surprisingly, the presentation of unnecessary text did not consistently hamper learning about the process of mitosis in the two Experiments (i.e., only a small negative effect of separated unnecessary text in Experiment 1, but not in Experiment 2; no negative effect of integrated unnecessary text in both experiments). Because we did not reliably find an initial negative effect of the unnecessary text on learning, the question of whether task experience would reduce or eliminate that negative effect could not be answered. A possible explanation for why the unnecessary information did not initially have a negative effect on learning even though it was processed, might lie in the nature of the extraneous information, that is, in whether it is irrelevant or unnecessary. It is possible that the negative effects of irrelevant information on learning would be larger than the effects of unnecessary information. That is, processing irrelevant information not only takes up working memory capacity but might actively interfere with learning the essential information, by disrupting the processing, organization, and integration of essential information. Processing unnecessary information on the other hand (which is identical in content to the essential information), does take up working memory capacity but may interfere less with learning the essential information.

Another potential explanation might lie in the amount of time that learners had available. We based the time per slide on the average study time of eight participants in a pilot study, which can therefore be assumed to have been sufficient for most of our participants. It is possible that processing unnecessary information would start to hamper learning when there is time pressure. When there is little time available for processing, any time spent on the unnecessary information goes at the expense of thoroughly processing essential information, and as a result, learning is hampered. In the present study, learners may have had sufficient time for processing all sources of information, which would explain why their attention to the unnecessary text (as demonstrated in Experiment 2) did not significantly increase experienced cognitive load and did not negatively affect learning as measured by either the cloze tests or the picture test. Systematically varying the presentation time in future research could shed some light on this issue.

Limitations and Future Research

It is interesting that we replicated the finding that attention to extraneous information wanes with the present materials, as these are more ecologically valid and more complex than the word learning materials in the studies by Rop and colleagues (Rop, Van Wermeskerken et al., in press; Rop, Verkoeijen et al., 2017).

However, a possible limitation of the present study, which might perhaps also explain the lack of effects on learning outcomes, is that the different phases of the process of mitosis are not fully independent of each other. As each phase is building on the information that was provided in the previous phase, the processing of later slides might have been dependent on how well information from the previous slides had been learned. Moreover, some phases might be more complex than others, which is also suggested by the differences in processing time per slide. Another potential limitation is that the cloze test mostly tested retention of the essential text; it is possible that the results regarding learning outcomes would be different when the test would assess understanding (e.g. by means of inference questions).

Concluding, the results of this study are interesting in that they provide evidence that learners adapt their study strategy and start to ignore unnecessary information with increasing task experience. However, our results also call for further research aiming to pinpoint conditions under which extraneous information presentation negatively affects learning, and employing eye-tracking methodology to study the attention allocation processes during learning may help accomplish this (see also Van Gog & Scheiter, 2010). Next to the nature of the information (irrelevant vs. unnecessary), the format of the information (textual vs. pictorial) and the layout of the information (integrated vs. separated), the role of time on task and the complexity of the learning and test materials should be investigated.

Chapter 5

The Effect of Layout and Pacing on Learning from Diagrams with Unnecessary Text

This chapter has been submitted for publication as:

Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2017). *The Effect of Layout and Pacing on Learning from Diagrams with Unnecessary Text*. Manuscript submitted for publication.

Abstract

While the presentation of extraneous (i.e., irrelevant or unnecessary) information hinders learning, it is unclear whether and how layout and pacing influence this effect. In two experiments participants learned how the heart functions using four different lay-outs: a diagram presented without unnecessary text (diagram only); with unnecessary text separated from the diagram (separated) or integrated into the diagram (integrated); or with separated unnecessary text and the instruction to integrate (integration instruction). In Experiment 1 study time was self-paced for half of the participants, and system-paced for the other half. There were no effects of layout and of pacing on learning, although system-pacing was more effortful than self-pacing. In Experiment 2, which was system-paced and employed eye tracking, the integrated condition showed worse learning outcomes than the separated condition. Moreover, in the integrated condition participants made more integration attempts between the unnecessary text and the diagram than in the separated condition.

Introduction

The majority of contemporary instructional materials are multimedia materials, comprising a combination of text (written or spoken) and pictures (e.g., diagrams, illustrations, graphs, or animations). Learning from multimedia has been widely investigated in research inspired by the Cognitive Theory of Multimedia Learning (Mayer, 2014) and Cognitive Load Theory (Sweller, Ayres, & Kalyuga, 2011). This research has shown that when a mutually referring text and picture are unintelligible in isolation, the combination of text and picture tends to improve learning (the multimedia effect, Butcher, 2014; Mayer, Bove, Bryman, Mars, & Tapangco, 1996). However, when either text or picture are intelligible in isolation, their combination might not help, and could even hamper learning (the redundancy effect, Kalyuga & Sweller, 2014; the coherence effect; Mayer & Fiorella, 2014). Accordingly, the redundancy and coherence principles of multimedia learning state that presentation of extraneous information should be avoided, because it hampers learning compared to instructional materials in which this information is eliminated.

Why Does Extraneous Information Hamper Learning?

The negative effect of extraneous information on learning can be explained in terms of working memory load. Working memory is limited in both capacity and duration (e.g., Baddeley, 2000; Barrouillet & Camos, 2007; Cowan, 2001; Miller, 1956). Regarding the limited capacity, Miller (1956) proposed that our memory span is limited to 'seven plus or minus two' chunks, where a chunk is a collection of items that are remembered together, although more recent research suggests this number is closer to four (Cowan, 2001). Regarding the limited duration, Barrouillet and Camos (2007) stress the importance of a time limit on working memory. They propose that working memory resources have to be shared between maintenance of 'old' information (prior knowledge from long-term memory or previously processed information during task performance), and processing of new, incoming information. The higher the number of old information elements that have to be maintained and the faster new elements need to be processed, the higher the working memory load. Therefore, a task that is cognitively undemanding when time is unlimited can become very demanding when time is limited.

Learning (i.e., schema construction/elaboration/automation in long-term memory; Sweller, 1994) requires that old information is maintained in working memory and successfully integrated with the new information that the learner selects (by attending to it) from the learning materials. When the selection, organization, or integration processes are disrupted, learning is hampered (Mayer, 2014). Extraneous information that is added to instructional materials can hamper learning when learners attend to, process, and attempt to integrate the extraneous information with the essential information, as this extraneous processing depletes valuable working memory resources. Consequently, these working memory resources cannot be allocated to processing the essential information that is relevant for learning.

The negative effects of extraneous information processing on learning have been shown with a wide variety of materials (for reviews, see Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). A seminal study regarding the negative effect of extraneous information on learning was performed by Chandler and Sweller (1991). They studied the relationship between extraneous text added to diagrams and learning outcomes, in a series of six experiments. In all experiments, study time was self-paced. In Experiment 1-4, the learning materials were in the domain of electrical engineering, while Experiments 5 and 6 used learning materials about the human circulatory system. In Experiment 1 and 6, the text contained essential information not conveyed in the diagram. In this case, participants performed better when the text was integrated in the diagram (integrated condition) than when it was presented spatially separated (separated condition; this is evidence of the 'split-attention effect'; see Ayres & Sweller, 2014). In Experiment 2, the text was unnecessary for learning, as the information it described could also be inferred from the diagram. In this case, participants showed better learning outcomes in the separated condition than in the integrated condition, presumably because the unnecessary information was harder to ignore in the integrated condition. In Experiment 3, all participants learned with the separated learning materials used in Experiment 2, but half of the participants were instructed to study the diagram, while the other half was instructed to read and integrate the textual information with the diagram (integration instruction condition). Learning outcomes were lower in the latter group, presumably because –in line with Experiment 2- participants in the integration instruction condition spent cognitive resources on processing the unnecessary text. In Experiment 4 and 5 participants learned in one of three conditions: diagram only, integrated, and integration instruction. Learning outcomes were significantly higher in the diagram-only condition than in the integrated and integration instruction conditions. Although the integrated condition performed numerically better than the integration instruction condition, this difference was not significant.

These results show that the presentation of unnecessary information alongside a diagram hampers learning in a self-paced learning environment. This effect seems to be stronger when the unnecessary information comes integrated with the essential information or when learners are instructed to integrate it, than when the information is presented spatially separated. It seems that in a separated layout, learners may be able to ignore the unnecessary information, at least under self-paced conditions.

The Effect of Layout and Pacing on Extraneous Processing

The Chandler and Sweller (1991) study shows that the layout of the materials influences the occurrence and strength of the negative effect of extraneous information on learning. Other studies support this finding. For example, Oksa, Kalyuga, and Chandler (2010) provided evidence that integrated extraneous information in the form of explanatory notes hampered interpretation of Shakespearian texts as a function of prior knowledge. While integrated explanatory notes improved text comprehension for low prior knowledge learners, for whom the notes were essential, they hampered comprehension for experts on Shakespearian texts, who did not need the notes to comprehend the text. The authors surmised that the experts were unable to ignore the explanatory notes, and that the extraneous processing disrupted their reading process.

Interestingly, Chandler and Sweller's (1991) findings from Experiments 2 and 3 showed that unnecessary text that is presented spatially separated from the diagram, led to better learning than an integrated text and diagram or instructions to integrate text and diagram. This suggests that learning was not hampered, or not hampered to the same extent, by unnecessary text when it was presented separately. An interesting question therefore, is whether students in the separated condition would be able to ignore spatially separated unnecessary text to such an extent that it does not hamper their learning compared to a diagram-only condition. This question cannot be answered based on the Chandler and Sweller study because they did not make a direct comparison of all four (i.e. diagram only, separated, integrated, integration instruction) conditions: Experiments 2 and 3 did not include a diagram-only condition, and Experiments 4 and 5 did not include a separated condition *without* the integration instruction.

This question of whether learners would be able to ignore unnecessary information when it is presented spatially separated from the essential information is relevant, because recent research suggests that students may learn to ignore extraneous information once they have gained some task experience (Rop, Van Wermeskerken, De Nooijer, Verkoeijen, & Van Gog, in press; Rop, Verkoeijen, & Van Gog, 2017). However, in these studies by Rop and colleagues, the extraneous information was pictorial, not textual, and mismatched the verbal information that participants had to remember. In the materials from Chandler and Sweller's (1991) study, the extraneous information is textual, which may be harder to ignore as learners often focus more quickly and more strongly on text than on the associated pictures (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010). Thus, it is safe to assume that unnecessary text is harder to ignore than pictures. Moreover, in the Chandler and Sweller study, the extraneous information was unnecessary, in the sense that the information provided by the text could also be inferred from the diagram, but it was not irrelevant or mismatching (i.e., it did have a relation with the picture). As such, it may be harder for learners to come to realize that they can ignore the text. Therefore, the present study addresses the question of whether learners would be able to ignore unnecessary textual information when it is presented spatially separated from the diagram, by building on the Chandler and Sweller study and comparing a diagram only, separated layout, integrated layout, and separated layout with integration instruction condition.

In addition, we study the effects of pacing in the present study. In the Chandler and Sweller (1991) study, learning was self-paced. Consequently, the time that participants spent on processing the unnecessary text did not have to go at the expense of time spent on processing the essential information from the diagram. It is likely that when learning is system-paced, the negative effect of processing unnecessary information on learning would be even stronger. That is, when the available time is limited, any extraneous information processing goes directly at the expense of essential information processing (cf. Barrouillet & Camos, 2007), which should lead to a stronger negative effect of extraneous information on cognitive load (i.e., higher) and learning outcomes (i.e., lower). Kalyuga, Chandler, and Sweller (2004; see also Mayer & Jackson, 2005) provide some empirical evidence for this hypothesis. In two experiments, one system-paced and one self-paced, participants had to learn how to use a cutting speed nomogram with or without unnecessary text.

While both experiments showed a negative effect of the unnecessary text, the effect sizes were larger in the system-paced experiment. However, this study did not directly manipulate learning in a self-paced environment vs. learning in a system-paced environment within one experiment, so a causal relationship between pacing and the negative effect of unnecessary text on learning could not be established.

The Present Study

The present study aimed to replicate and extend the results from the seminal study of Chandler and Sweller (1991) by making a direct comparison of all four conditions (i.e., diagram only, separated, integrated, and integration instruction) and by investigating the effects of self- vs. system-pacing, using (an adapted version of) their materials about the human circulatory system (Experiment 5). We conducted two experiments in which participants learned about the human circulatory system using a self-contained diagram of the heart, which was accompanied by unnecessary textual explanations. The text was unnecessary for learning because it described processes that could also be inferred from the diagram, which, according to the findings by Chandler and Sweller (1991), could be understood without references to the text. In both experiments, four conditions were compared: The unnecessary text was not presented at all (diagram-only condition), or was presented separated from the diagram (separated condition), integrated in the diagram (integrated condition), or separated from the diagram with the instruction to integrate it (integration instruction condition).

In Experiment 1, half of the participants learned at their own pace (self-paced conditions) while learning was time constrained for the other half of the participants (system-paced conditions). Learning outcomes were assessed using three retention tests (out of the six tests used by Chandler and Sweller, 1991). With regard to learning outcomes, we expected to replicate Chandler and Sweller's overall findings (i.e., 1: diagram only > integrated & integration instruction; 2: separated > integrated & integration instruction) and sought to answer the open question of whether students are able to ignore spatially separated unnecessary text to such an extent that it does not hamper their learning compared to a diagram only condition (i.e., diagram only = separated > integrated & integration instruction), or whether the presentation of unnecessary text should be avoided entirely as it does hamper learning even when presented spatially separated (i.e., diagram only > separated > integrated & integration instruction). In addition, we hypothesized that system-pacing would aggravate the negative effects of unnecessary text on learning particularly in the integrated condition and the integration-instruction condition because these conditions are assumed to impose the highest cognitive load on the learner. Experienced cognitive load was measured by having participants rate their invested mental effort (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). With regard to cognitive load experienced while studying, we expected the reverse of the learning outcomes pattern (i.e., diagram only ≤ separated < integrated & integration instruction) and pacing (i.e., integrated & integration instruction conditions: self-paced < system-paced).

In Experiment 2, we employed eye tracking to investigate how the presentation of unnecessary text affects attention allocation under system-paced conditions. This allowed us to directly measure whether or not learners paid less attention to the unnecessary text when it was spatially separated from the diagram compared to integrated in the diagram, and compared to when they had been given integration instructions.

Experiment 1

Method

Participants

Participants (*n* = 302) were recruited via Amazon's Mechanical Turk (Paolacci, Chandler, & Ipeirotis, 2010), and were paid 0.60 dollar for their participation (which took about six to seven minutes). A priori we decided that participants would be excluded when they used a device with a small screen and could not see the diagram and explanations sufficiently (i.e., could not see at least six lines of explanations) without scrolling (*n* = 61); when they participated in the experiment twice (*n* = 5); or when they participated in a noisy environment (i.e., self-reported noise of seven or higher on a scale of one to nine; *n* = 5). Furthermore, we decided post-hoc to exclude participants who experienced technical problems (in the system-paced condition, for some participants the diagram was presented for longer than the pre-defined presentation time; *n* = 3); who did not follow orders (*n* = 2); and who reported using memory influencing drugs (*n* = 1). Thus, our final sample comprised 225 participants (M_{age} = 37.40 years, *SD* = 12.97, range 18-87; 108 females).

Design

Participants were randomly assigned to one of the eight conditions resulting from a 2 x 4 factorial design with between-subjects factors 'pacing' (self-paced vs. system-paced) and 'layout' (diagram only, separated, integrated, integration instruction): system-paced diagram only (*n* = 35), system-paced separated (*n* = 21), system-paced integrated (n = 28), system-paced integration instruction condition (n = 26); and self-paced diagram only (n = 34), self-paced separated (n = 37), self-paced integrated (n = 24), and self-paced integration instruction (n = 30)[1].

Materials.

The materials were programmed and presented in Qualtrics software (Qualtrics, Provo, UT).

Subjective and objective prior knowledge measures. In order to measure participants' prior knowledge on this topic, a subjective and an objective measure were used. The subjective measure asked participants to indicate how much they knew about the blood flow in our body, ranging from one (*nothing at all*) to five (*a great deal*). The objective measure asked participants to name the four major components of the heart and the two arteries through which blood exits the heart.

Learning material. The learning material consisted of a simplified diagram of the circulatory system, which was adapted from Chandler and Sweller (1991, Experiment 5) with slight modifications. We vectored the image, added color (blue for the deoxygenated blood and red for the oxygenated blood), changed the font size and type, and increased the resolution of the image. The diagram was either the sole source of information on the screen (diagram-only condition, see Figure 1), or it was accompanied by unnecessary textual explanations separated from the diagram (separated and integration instruction conditions, see Figure 1), or it was accompanied by the same explanations integrated in the diagram (integrated condition, see Figure 1). The unnecessary text consisted of 96 words. In the diagram only, integrated, and separated conditions, participants were instructed as follows: "Please do your best studying the diagram, because afterwards you will be tested on your knowledge of the blood flow in the heart and body". The instruction for the participants in the integration instruction condition was: "Please try to integrate the diagram and text as much as possible, because afterwards you will be tested on your knowledge of the blood flow in the heart and body".

[1] This is the n after exclusion (hence the unequal sample sizes); see participants section

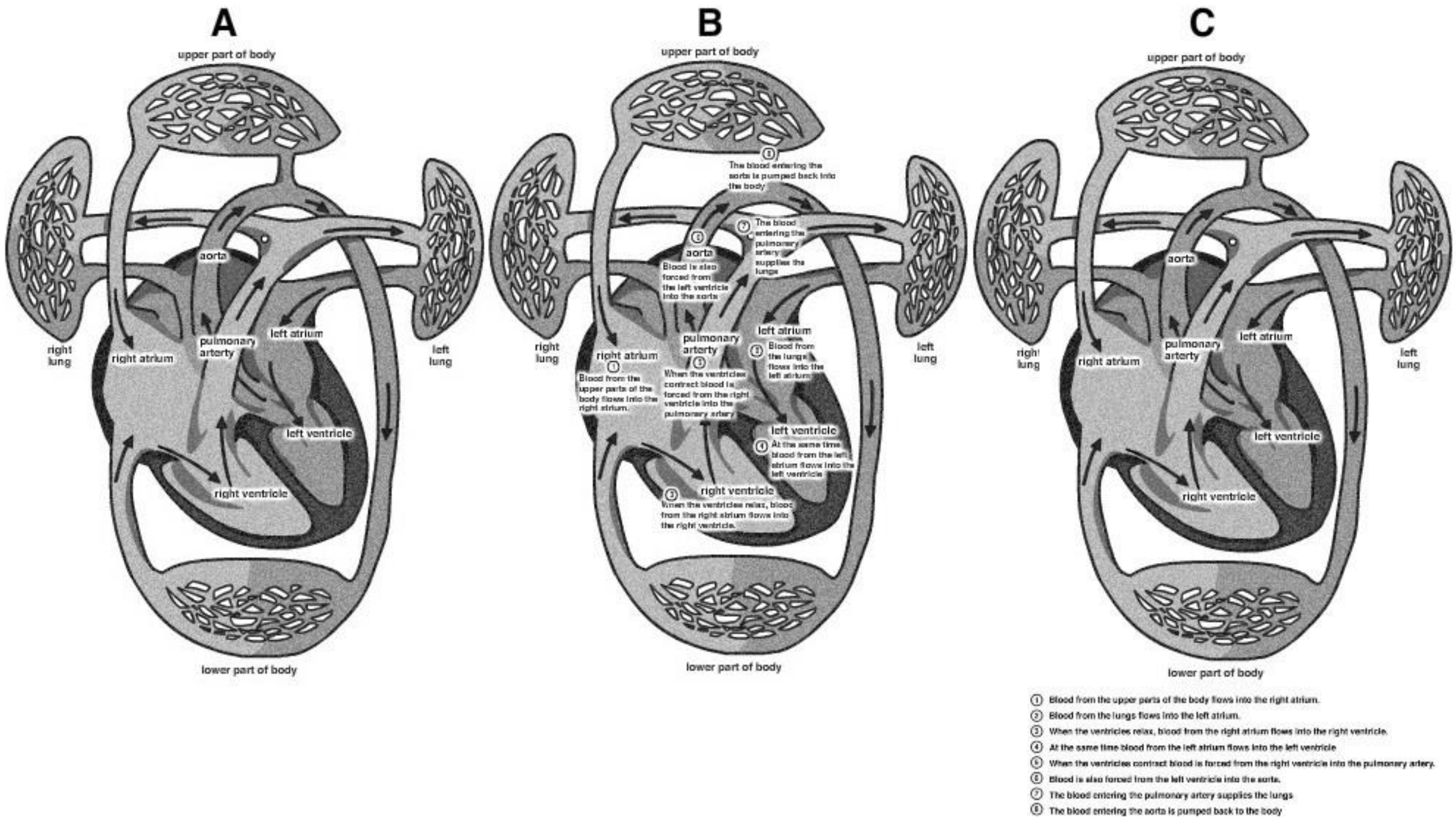


Figure 1. Learning materials in Experiment 1 for the diagram-only condition (1A), the integrated condition (1B), and the separated and integration instruction conditions (1C).

Participants in the self-paced conditions could spend as much time studying the learning material as they wished, while their time on task was measured. Participants in the system-paced conditions had 80s to study the learning material, which was ca. 10s more than the average time spent by participants in the diagram-only condition in the Chandler and Sweller study (1991, i.e., 69.1s). Participants were informed about this time limit.

Posttest. The posttest assessed retention of the learning materials using three of the six outcome measures used by Chandler and Sweller (1991). These were the measures most suitable for online, computer-based testing, as the other three outcome measures required physical manipulation of the diagram (e.g., placing arrowheads to indicate the correct flow of the blood). In the Chandler and Sweller study the first two outcome measures showed a significant advantage of diagram-only presentation while this advantage was marginally significant for the third. First, participants were asked to name the four major components of the heart and the two arteries through which blood exits the heart (components test). This was a single question for which participants had to give up-to six answers. Then, they were asked to complete two blood flow chains presented on separate slides (blood chains test): Participants were told that the second component was either the left atrium or the right atrium (e.g., __, left atrium, __, __, __) and had to complete the first, third, fourth, and fifth component. Finally, participants were given five fill-in-the-gap questions, all presented on separated slides, in the form of 'blood in the left ventricle flows to the __' (blood flow test). The five questions regarded the left atrium, right ventricle, left ventricle, pulmonary artery, and aorta. The three outcome measures were always presented in the same order, while the order of the questions within each measure was randomized. The posttest was self-paced, and participants could not go back to previous questions to adjust their answers.

Mental effort. Participants were asked to indicate how much effort they invested in studying the learning material and in answering the posttest questions on a nine point rating scale (Paas, 1992), ranging from one (*extremely low effort*) to nine (*extremely high effort*). Mental effort is an indicator of experienced cognitive load (see Paas et al., 2003).

Control questions and demographic questionnaire. To obtain information on the circumstances under which the experiment was completed, participants were asked some control questions after the posttest. Participants in the separated and integration instruction conditions were asked whether they could see the entire diagram and all explanations without scrolling on a scale of one (*yes*) to nine (*no, I could not see any of the explanations*). Participants in the diagram-only and integrated conditions were asked whether they could see the entire diagram without scrolling with a yes or no question. All participants had to self-report on the noise in their environment on a scale ranging from one (*quiet and no distractions*) to nine (*noise and many distractions*). Finally, participants were asked to provide some demographic information, namely their age, gender, and highest achieved education level.

Procedure

Participants first received a short introduction about the experiment, and instructions on what they had to do. Then, they filled out the subjective prior knowledge questions and completed the objective prior knowledge test, after which they studied the learning material, with layout and pacing depending on assigned condition. After the learning phase, participants were asked to indicate how much mental effort they had invested in studying the learning material. Then, participants completed the posttest and reported how much mental effort they invested in answering the posttest questions. Finally, they answered the control and demographic questions.

Scoring

For all measures, participants were awarded 1 point per correct response. When they gave a partially correct answer (e.g., pulmonary instead of pulmonary artery), they received 0.5 points. If they did not provide an answer, or if it was completely wrong, 0 points were awarded. Thus, participants could get a maximum of six points for the component test, eight points for the blood chains test (four for the left, and four for the right chain), and five points for the blood flow test. A random subset of the data (10.67%) was scored by a second rater, and interrater reliability was high ($\kappa = .90$).

Results

In this study, we maintained an alpha level of .05. For the parametric tests, effect size measures used were partial eta-squared (η_p^2) and Cohen's *d*. Both can be interpreted in terms of small ($\eta_p^2 \sim .01$, $d \sim 0.2$), medium ($\eta_p^2 \sim .06$, $d \sim 0.5$), and large ($\eta_p^2 \sim .14$, $d \sim 0.8$) effect sizes (Cohen, 1988).

Prior knowledge

To check whether conditions did not differ on prior knowledge (data presented in Table 1), we performed a 2 (pacing) x 4 (layout) ANOVA on the self-reported prior knowledge , which showed a main effect of pacing, $F(1, 217) = 7.92$, $p = .005$, $\eta_p^2 = .04$, indicating that participants in the self-paced conditions ($M = 2.46$, $SD = 0.83$) rated their domain knowledge higher than participants in the system-paced conditions ($M = 2.16$, $SD = 0.75$). There was no effect of layout and no interaction (Fs < 1). The objective prior knowledge test did not confirm higher prior knowledge in the self-paced conditions though. Because the assumptions of normality were violated for the objective domain knowledge test, we performed non-parametric tests on these data. A Mann-Whitney test indicated no significant differences between the self-paced and system-paced conditions, $U = 5552.50$, $p = .111$. Kruskal-Wallis tests showed no significant differences between layout conditions when the presentation was self-paced, $\chi^2(3) = 2.52$, $p = .472$, or system-paced, $\chi^2(3) = 1.18$, $p = .759$.

Table 1. *Mean (SD) Self-estimated Prior Knowledge (range = 1-5) and Mean (SD) and Median (Range) Performance on the Objective Domain Prior Knowledge Test (max = 6) as a Function of Pacing and Layout in Experiment 1.*

		Self-Estimated	Objective	
		Mean (SD)	Mean (SD)	Median (Range)
Self-paced	Diagram only	2.56 (0.93)	2.59 (2.16)	2.00 (6.00)
	Separated	2.59 (0.89)	2.09 (2.14)	1.50 (6.00)
	Integrated	2.42 (0.65)	2.67 (2.07)	2.50 (6.00)
	Integration instruction	2.27 (0.78)	2.70 (1.61)	3.00 (6.00)
	Total	**2.46 (0.83)**	**2.52 (1.99)**	**2.00 (6.00)**
System-paced	Diagram only	2.20 (0.80)	2.44 (2.20)	1.50 (6.00)
	Separated	2.10 (0.77)	2.02 (1.41)	2.00 (4.00)
	Integrated	2.21 (0.74)	1.82 (1.61)	1.50 (5.00)
	Integration instruction	2.12 (0.71)	1.90 (1.79)	1.00 (6.00)
	Total	**2.32 (0.80)**	**2.07 (1.81)**	**1.50 (6.00)**

Time on task

We checked how much time participants in the self-paced conditions spent on the learning materials: On average, participants in the self-paced condition studied the materials for 100.77s ($SD = 106.77$, $Mdn = 67.83$). Mean time on task was highest in the self-paced diagram-only condition (129.25s, $SD = 147.69$; $Mdn = 71.25$), followed by the self-paced integrated condition (104.06s, *SD* = 107.62; *Mdn* = 68.24), the self-paced integration instruction condition (98.49, *SD* = 69.87; *Mdn* = 84.43), and the self-paced separated condition (64.42s, *SD* = 64.42; *Mdn* = 42.27) Thus, except for the separated condition, the average time on task was higher in the self-paced than in the system-paced (i.e., 80s) conditions

Learning outcomes

For all learning outcome measures (see Table 2), the assumptions of normality were violated. Therefore, we performed non-parametric tests on these data. On the components test, a Mann-Whitney test showed no significant differences between the self-paced and system-paced conditions, $U = 6069.00$, $p = .574$. Kruskal-Wallis tests showed no significant differences between layout conditions when the presentation was either self-paced, $\chi^2(3) = 0.47$, $p = .926$, or system-paced, $\chi^2(3) = 4.40$, $p = .221$.

Table 2. *Mean (SD) and Median (Range) Performance on the Components (max. = 6), Blood Chains (max. = 8), and Blood Flow (max. = 5) Tests as a Function of Pacing and Layout in Experiment 1.*

		Components		Blood chains		Blood flow	
		Mean (SD)	Median (Range)	Mean (SD)	Median (Range)	Mean (SD)	Median (Range)
Self-paced	Diagram only	4.81 (1.70)	5.75 (6.00)	2.32 (2.40)	2.00 (8.00)	2.65 (2.66)	3.00 (5.00)
	Separated	5.00 (1.59)	6.00 (5.00)	2.41 (2.66)	1.50 (8.00)	3.02 (1.74)	3.50 (5.00)
	Integrated	5.02 (1.28)	5.75 (6.00)	3.54 (2.68)	3.00 (8.00)	2.92 (1.74)	3.00 (5.00)
	Integration instruction	4.87 (1.68)	6.00 (5.00)	3.07 (2.98)	2.00 (8.00)	2.58 (1.81)	2.75 (5.00)
	Total	**4.91 (1.57)**	**6.00 (6.00)**	**2.79 (2.69)**	**2.00 (8.00)**	**2.77 (1.70)**	**3.00 (5.00)**
System-paced	Diagram only	5.01 (1.29)	5.50 (5.50)	1.56 (1.92)	1.00 (6.50)	2.47 (1.70)	2.00 (5.00)
	Separated	4.67 (1.64)	5.00 (6.00)	1.86 (1.80)	2.00 (7.00)	1.90 (1.68)	1.50 (5.00)
	Integrated	4.95 (1.13)	5.00 (3.50)	2.34 (2.26)	1.50 (7.00)	2.36 (1.31)	2.00 (5.00)
	Integration instruction	5.35 (1.21)	6.00 (4.00)	2.60 (2.44)	2.00 (8.00)	2.50 (1.57)	2.75 (5.00)
	Total	**5.01 (1.31)**	**5.50 (6.00)**	**2.06 (2.13)**	**1.00 (8.00)**	**2.34 (1.57)**	**2.00 (5.00)**

On the blood chains test, the self-paced conditions seemed to outperform the system-paced conditions, but a Mann-Whitney test showed that this difference was not significant, $U = 5471.50$, $p = .075$. Kruskal-Wallis tests indicated no performance differences between layout conditions when the presentation was either self-paced, χ^2 (3) = 3.65, $p = .302$, or system-paced, χ^2 (3) = 4.13, $p = .248$.

Finally, on the blood flow test, the Mann-Whitney test showed a marginally significant advantage for the self-paced conditions compared to the system-paced conditions, $U = 5393.00$, $p = .054$. Again, however, Kruskal-Wallis tests indicated no performance differences between layout conditions when the presentation was either self-paced, χ^2 (3) = 1.42, $p = .702$, or system-paced, χ^2 (3) = 2.27, $p = .519$.

Invested Mental Effort.

The data regarding invested mental effort can be found in Table 3. We performed a 2 (pacing) x 4 (layout) ANOVA on the invested mental effort during learning, which showed a main effect of pacing, $F(1, 217) = 5.37$, $p = .021$, $\eta_p^2 = .02$, indicating that participants in the system-paced conditions invested more mental effort during learning than participants in the self-paced conditions. We found no effect of layout, $F(3, 217) = 2.04$, $p = .109$, $\eta_p^2 = .03$, and no interaction, $F(3, 217) = 1.11$, $p = .347$, $\eta_p^2 = .02$. The same ANOVA on mental effort invested during the posttest indicated no effect of pacing, $F(1, 217) = 1.20$, $p = .275$, $\eta_p^2 = .01$, layout, $F < 1$, and no interaction $F < 1$.

Table 3.

Table 3: *Mean (SD) Invested Mental Effort (max. = 9) During the Learning Phase and Posttest as a Function of Pacing and Layout in Experiment 1.*

		Learning Phase	Posttest
Self-paced	Diagram only	7.50 (1.31)	7.76 (1.28)
	Separated	6.67 (1.82)	7.41 (1.37)
	Integrated	6.71 (1.85)	7.25 (1.92)
	Integration instruction	7.27 (1.08)	7.30 (1.37)
	Total	**7.08 (1.54)**	**7.45 (1.47)**
System-paced	Diagram only	7.49 (0.95)	7.63 (1.09)
	Separated	7.19 (1.12)	7.43 (1.29)
	Integrated	7.57 (1.14)	7.75 (1.29)
	Integration instruction	7.58 (1.36)	7.73 (1.46)
	Total	**7.47 (1.13)**	**7.65 (1.26)**

Discussion

In contrast to our hypotheses regarding layout, we did not replicate Chandler and Sweller's (1991) findings that learning outcomes would be significantly lower in the integrated and integration instruction conditions than in the diagram-only condition or the separated condition. With regard to our question of whether the separated condition would perform as well or worse than the diagram-only condition, we found no difference among those conditions. However, given the failure to replicate their respective superiority over the integrated and integration instruction conditions, that finding cannot be meaningfully interpreted.

Regarding pacing, we found that self-paced learning was less effortful than system-paced learning. In other words, in line with our hypothesis, system-pacing did indeed increase perceived cognitive load. Moreover, self-pacing led to a numerical (but not statistically significant, p = .075) advantage in performance on the blood chains test and a marginally significant (p = .054) advantage in performance on the blood flow test compared to system-pacing. However, given the lack of differences among layout conditions in either self-paced or system-paced learning, we found no evidence that system-pacing would aggravate the negative effects of presenting unnecessary text on learning.

We see two potential explanations for our failure to replicate Chandler and Sweller's (1991) findings. First, to control for possible differences in prior knowledge, we added an objective prior knowledge test, which Chandler and Sweller (1991) did not have in their experiment. This test might have guided participants' learning, which could have neutralized any effects of the unnecessary text. Indeed, research into test-potentiated learning shows that attempting to retrieve information may improve later encoding, even when the initial retrieval attempt was unsuccessful (Arnold & McDermott, 2013; Little & Bjork, 2016). We cannot rule out that this could have led to improved learning performance in the conditions with unnecessary text, in which case we should see negative effects of unnecessary text on learning when not giving a pretest. Second, we ran our experiment on Mechanical Turk, and although Mechanical Turk can yield results similar to lab studies (Paolacci et al., 2010), this type of online testing reduces experimental control somewhat. For instance, the time on task data from the self-paced conditions showed substantial variance, and we cannot be sure that participants spent all this time studying the learning materials (i.e., they might have been interrupted or attending to another task). However, while this may at least partly explain the large variance in learning outcomes, it does not explain the absence of mean/median condition differences.

Therefore, we dropped the objective prior knowledge test in Experiment 2 and conducted this experiment in a lab context. We employed eye-tracking to gain more insight into the question of how the presentation of unnecessary text affects attention allocation during learning, as it was unclear to what extent students processed the unnecessary information in Experiment 1.

Experiment 2

In Experiment 2, we compared the four system-paced conditions (i.e., diagram only, separated, integrated, integration instruction) on learning outcomes and perceived cognitive load in a more controlled lab setting, and excluded the objective prior knowledge test for the reason mentioned in the Discussion section of Experiment 1. We employed eye-tracking to directly measure to what extent students processed the unnecessary text, and to address our assumption that participants in the separated condition would attend less to the unnecessary text than participants in the integrated and the integration instruction conditions. This should result in shorter fixation duration on the unnecessary text and less transitions between unnecessary text and the relevant parts of the diagram. With regard to processing the relevant information, that is, the arrows and labels in the diagram, we would expect that participants in the diagram-only condition would pay most attention to this information, followed by the separated condition and the integrated and integration instruction conditions, with the question being whether the diagram only and separated condition would differ from each other.

Method

Participants and design.

Participants were 132 German University students (M_{age} = 21.15 years, *SD* = 2.34; 106 female) who participated for course credit or 3 Euro. All participants had normal or corrected-to-normal vision. One participant indicated after completing the experiment that he/she wanted to retract his/her data. Moreover, nine participants were removed from the analyses they had accidentally ended the learning phase by pressing the spacebar, which was not allowed. The remaining 122 participants were distributed across the four conditions (to which they had been randomly assigned) as follows: diagram only (*n* = 29), separated (*n* = 33), integrated (*n* = 32), and integration instruction (*n* = 28).

Materials

The materials were programmed and presented in SMI Experiment Center (Version 3.6; SensoMotoric Instruments). They were largely similar to Experiment 1, with two exceptions: 1) the textual explanations in the separated and integration instruction conditions were presented next to the diagram rather than underneath it (see Figure 1 and 2) and 2) the materials were presented system-paced at 100s (mean time on task in the self-paced conditions in Experiment 1) to ensure that participants had sufficient time. Furthermore, due to a programming error, the answers for the components test were not saved in Experiment 2.

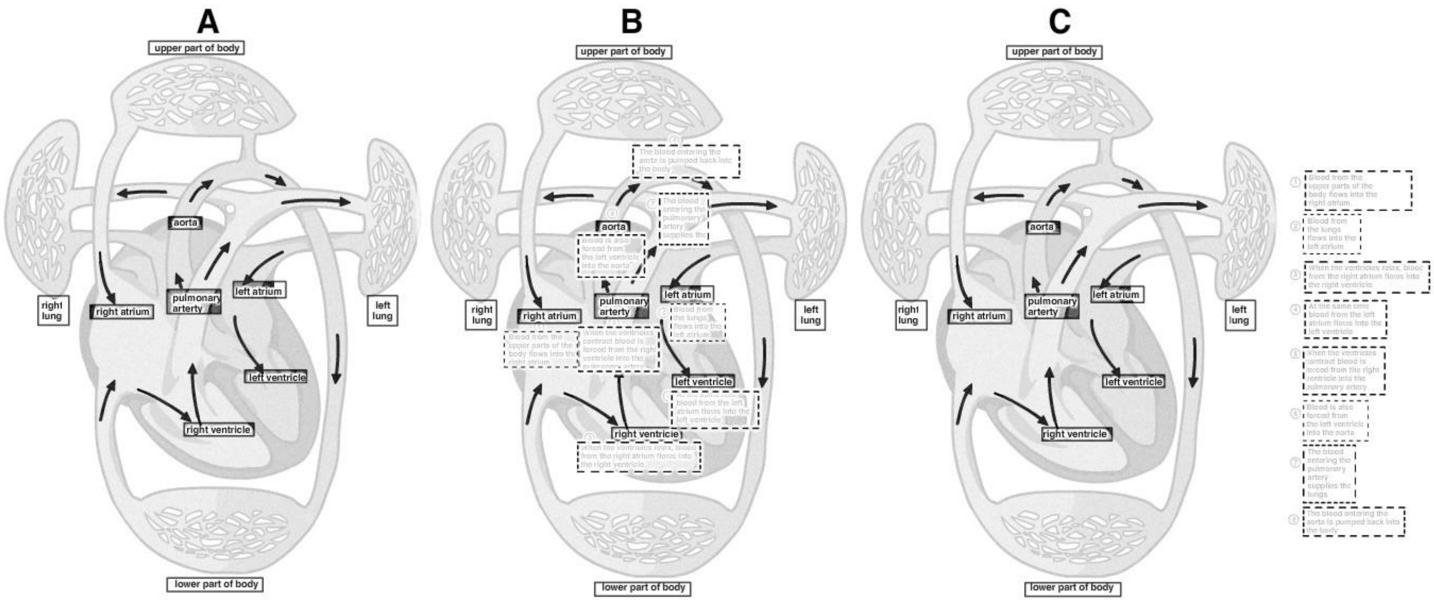


Figure 2. Learning materials in Experiment 2 for the diagram-only condition (2A), the integrated condition (2B), and the separated and integration instruction conditions (2C). The different AoIs are presented in a box with a continuous line (relevant text), a box with a dotted line (unnecessary text), and in black (arrows).

Apparatus

The materials were presented in SMI Experiment Center (Version 3.6; SensoMotoric Instruments) on a monitor with a resolution of 1920 x 1080 pixels. Participants' eye movements were recorded using SMI RED 250 Mobile eye trackers (SensoMotoric Instruments) that recorded binocularly at 250 Hz using SMI iView software (Version 4.2; SensoMotoric Instruments). SMI BeGaze software (Version 3.7; SensoMotoric Instruments) was used for data analysis.

Procedure

Participants were tested individually, or in groups of up to three participants simultaneously. At the start of the experiment, participants were seated in front of a laptop with the mobile eye tracker placed underneath, with their head approximately 60 cm from the screen. First, participants filled in their age and gender. After a short introduction about the experimental procedure, participants provided the subjective prior knowledge rating. Subsequently, the eye tracker was calibrated using a thirteen-point calibration plus four-point validation procedure, and participants were instructed to move as little as possible. Then the learning phase started, and participants were given 100 seconds to study the learning material, after which they rated how much mental effort they had invested in studying the material. Finally, participants completed the posttest and were asked to indicate how much mental effort they invested in answering the posttest questions. The experiment took 10-15 minutes to complete.

Data analysis

Posttest performance on the blood chain and blood flow tests was scored cf. Experiment 1. For the eye tracking analyses, we first checked the calibration accuracy and tracking ratio (i.e., the percentage of time for which the eye tracker actually measured the eye movements). We excluded two participants from the eye-movement data analyses (one from the integrated, and one from the integration instruction condition) because of inaccurate calibration (i.e., deviation from the four validation points exceeded 1°), and 28 additional participants (diagram-only: $n = 8$; separated: $n = 6$; integrated: $n = 10$; integration instruction: $n = 4$) because their tracking ratio was below 70%. The final sample for the eye-movement data analyses ($n = 92$) had an average tracking ratio of 88.52% ($SD = 6.51\%$), with a mean calibration accuracy of 0.27° ($SD = 0.13^{\circ}$), and was distributed across the conditions as follows: diagram-only ($n = 21$), separated ($n = 27$), integrated ($n = 21$), and integration instruction ($n = 23$).

Fixations were defined using a 40 °/s velocity threshold and a minimal duration of 100 ms (cf. Holmqvist et al., 2011). Areas of Interest (AoIs) were created on the diagram, covering the relevant text labels and the arrows indicating the blood flow (see Figure 2). In the separated, integrated, and integration instruction conditions, AoIs were also created on the unnecessary text lines (see Figure 2). This allowed us to compute the fixation duration on the unnecessary and essential AoIs; however, to control for differences in tracking ratio between participants, we calculated a relative measure of fixation duration by dividing the fixation duration on each AoI by the total fixation duration on the learning material (i.e., the sum of fixation duration on the different AoIs and white space).

To measure integration of the different sources of information (i.e., relevant text labels, arrows, unnecessary text), we used transitions between the different AoIs. We defined two types of transitions: unnecessary-relevant and relevant-relevant transitions. Unnecessary-relevant transitions are transitions between the unnecessary text and the relevant parts of the diagram (i.e., the relevant text labels and the arrows) or vice versa. Relevant-relevant transitions are transitions between and within two different relevant parts of the diagram, that is, between the relevant text labels and the arrows (or vice versa), within the arrows, and within the relevant text labels. To control for differences in tracking ratio we again divided the number of transitions by the total fixation duration on the learning material.

Results

Prior knowledge

A one-way ANOVA with layout as between-factor showed no differences between the diagram only (M = 2.03, SD = 0.73), integrated (M = 2.16, SD = 0.63), separated (M = 2.03, SD = 0.64), and integration instruction conditions (M = 1.96, SD = 0.43) on the subjective ratings of prior knowledge, $F < 1$.

Learning outcomes

For the two learning outcome measures, the assumptions of normality were violated and therefore non-parametric tests were conducted (see Table 4). On the blood chains test, a Kruskal-Wallis test showed a significant difference between conditions, χ^2 (3) = 14.97, p = .002. Follow-up Mann-Whitney tests with a Bonferroni corrected p-value (i.e., multiplying the p-value by six, the number of tests that were performed) showed that participants in the integrated condition had lower performance than participants in the separated condition, U = 292.00, p = .012 and participants in the integration instruction condition *U* = 222.00, *p* = .006. No other comparisons were significant, *p* > .396.

On the blood flow test a Kruskal-Wallis test showed significant differences between the conditions, χ^2 (3) = 12.91, *p* = .005. Follow-up Mann-Whitney tests (with Bonferroni corrected *p*-values) showed that participants in the integrated condition had lower performance than participants in the separated condition, *U* = 311.00, *p* = .024. None of the other comparisons yielded significant results, *p* > .060.

Table 4. *Mean (SD) and Median (Range) Performance on the Blood Chains (max. = 8) and Blood Flow (max. = 5) Tests as a Function of Layout in Experiment 2*

	Blood chains		Blood flow	
	Mean (SD)	Median (Range)	Mean (SD)	Median (Range)
Diagram only	2.00 (2.36)	1.00 (7.50)	1.28 (1.61)	0.50 (5.00)
Separated	2.80 (2.36)	2.50 (7.50)	2.06 (1.36)	2.00 (4.50)
Integrated	1.17 (1.88)	0.25 (8.00)	1.14 (1.12)	1.00 (5.00)
Integration instruction	2.64 (2.02)	2.25 (7.50)	1.82 (1.27)	2.00 (5.00)
Total	**2.15 (2.24)**	**1.50 (8.00)**	**1.58 (1.38)**	**1.00 (5.00)**

Invested Mental Effort

Data on the self-reported mental effort investment during the learning phase and on the posttest is reported in Table 5. A one-way ANOVA showed no significant differences among conditions in self-reported mental effort invested during learning, $F(3, 118) = 2.23$, $p = .089$, $\eta_p^2 = .05$, or during the posttest, $F < 1$.

Table 5: *Mean (SD) Invested Mental Effort (max. = 9) During the Learning Phase and Posttest as a Function of Layout in Experiment 2.*

	Learning Phase	Posttest
Diagram only	6.38 (1.08)	6.24 (1.35)
Separated	7.09 (1.13)	6.09 (1.72)
Integrated	6.59 (1.36)	5.91 (1.63)
Integration instruction	6.89 (1.07)	6.29 (1.61)
Total	**6.75 (1.19)**	**6.12 (1.58)**

Fixation Duration

See Table 6 for the fixation duration on the relevant information (i.e., text labels and arrows) and the unnecessary text. A one-way ANOVA on the fixation duration on the unnecessary text in the separated, integrated, and integration instruction conditions, showed no significant differences among conditions, $F(2, 71) = 2.46$, $p = .093$, $\eta_p^2 = .07$. For the fixation duration on the relevant information a one-way ANOVA revealed a significant main effect of layout, $F(3, 92) = 30.61$, $p < .001$, $\eta_p^2 = .51$. Bonferroni-corrected follow-up tests showed that, as expected, participants in the diagram-only condition spent more time looking at the relevant information than participants in the separated condition, $p < .001$, $d = 1.91$, integrated condition, $p < .001$, $d = 1.72$, and integration instruction condition, $p < .001$, $d = 2.35$. All other comparisons were not significant, smallest $p = .219$.

Transitions

See Table 6 for the unnecessary-relevant and relevant-relevant transitions. A one-way ANOVA on the unnecessary-relevant transitions revealed a significant main effect of layout, $F(2, 71) = 39.92$, $p < .001$, $\eta_p^2 = .54$. Participants in the integrated condition made more transitions between the unnecessary text and the relevant parts of the diagram than participants in the separated condition, $p < .001$, $d = 1.89$, and the integration instruction conditions, $p < .001$, $d = 1.77$. The analysis revealed no differences between the separated and integration instruction condition, $p > .999$.

Analysis of the relevant-relevant transitions again revealed a significant main effect of layout, $F(3, 92) = 8.83$, $p < .001$, $\eta_p^2 = .23$. Bonferroni-corrected follow-up tests showed that participants in the diagram-only condition made more relevant-relevant transitions than participants in the integrated, $p = .002$, $d = 0.94$; separated, $p < .001$, $d = 1.21$; and integration instruction conditions, $p < .001$, $d = 1.39$. All other comparisons were not significant, $p > .999$.

Table 6. *Mean (and SD) Fixation Duration on, and Transitions between, the Relevant Information and Unnecessary Text as a Function of Layout.*

	Fixation Duration		Transitions	
	Relevant	Unnecessary	Relevant-relevant	Unnecessary-relevant
Diagram only	0.24 (0.08)		0.12 (0.07)	
Separated	0.10 (0.04)	0.36 (0.10)	0.05 (0.05)	0.03 (0.02)
Integrated	0.12 (0.04)	0.43 (0.13)	0.05 (0.07)	0.20 (0.13)
Integration instruction	0.09 (0.04)	0.41 (0.13)	0.04 (0.04)	0.04 (0.03)
Total	**0.13 (0.08)**	**0.40 (0.12)**	**0.06 (0.06)**	**0.09 (0.10)**

Discussion

In Experiment 2, we again found no evidence for an overall negative effect of unnecessary text on learning. In contrast to Experiment 1, we did find some effects of layout: on both outcome measures, performance was higher in the separated condition than in the integrated condition. The integration instruction did not seem to influence learning as no differences emerged between this condition and the separated condition without integration instruction.

The eye-tracking measures indicate that the unnecessary text attracted attention, at the expense of attention to the relevant text labels and arrows, as participants in the unnecessary-text conditions looked less at the relevant parts of the diagram (and transitioned less between relevant parts of the diagram) than participants in the diagram-only condition. Participants in the separated condition spent as much time studying the unnecessary text as participants in the integrated and integration instruction conditions, while participants in the integrated condition made more transitions between the unnecessary text and relevant parts of the diagram than participants in the separated and integration instruction conditions.

This finding provides direct evidence that spatially integrating text and diagram affects participants' study strategy, inducing them to integrate the two sources of information more. This is in line with explanations (Ayres & Sweller, 2014) for why an integrated layout is preferable over a separated layout when both the text and the diagram are necessary for learning (see also Holsanova, Holmberg, & Holmqvist, 2009) and with research showing that transitions decreased as the distance between two sources of information increased (Bauhoff, Huff & Schwan, 2012). However, when the information is unnecessary for learning, the increased integration seems to hinder learning compared to a separated layout (though not compared to a diagram only).

The location of the unnecessary information was changed (from underneath to besides the diagram) in Experiment 2 to make the materials more suitable for eye tracking research. This location change might have influenced how the unnecessary text was processed in the separated and integration instruction conditions. However, this is highly unlikely as participants in these two conditions had similar posttest performance in Experiment 1 and 2.

General Discussion

The present study aimed to address the effects of layout and study time pacing on the negative influence of extraneous information on learning. We set out to replicate and extend the results from Chandler and Sweller (1991). They showed that it is better to leave out unnecessary text (diagram only) than to present it integrated into the diagram or separated from the diagram with the instruction to integrate it. They also demonstrated that a separated presentation (without integration instruction) led to better learning outcomes than an integrated presentation of unnecessary text. Their findings raised the interesting question of whether students in the separated condition would be able to ignore spatially separated unnecessary text to such an extent that it does not hamper their learning compared to a diagram only condition. Another open question was whether pacing would aggravate the effects of unnecessary information on learning, mostly when the unnecessary text is spatially integrated or when students are instructed to integrate spatially separated text. Therefore, we directly compared all four layout conditions under either system-paced or self-paced conditions: diagram only, separated, integrated, and integration instruction.

Unfortunately, we did not replicate most of Chandler and Sweller's (1991) findings. In Experiment 1, which was conducted online, we found no significant differences in learning outcomes among layout conditions. Thus, the presentation of unnecessary text did not hamper learning about the human circulatory system compared to studying only the diagram of the heart, regardless of whether participants learned at their own pace or under system pacing. However, self-paced learning was less effortful than system-paced learning, and seemed to increase posttest performance (numerically; this was not statistically significant), which is in line with theoretical arguments favoring self-pacing over system-pacing (Barouillet & Camos, 2007) and empirical studies (Kalyuga et al., 2004; Mayer & Jackson, 2005).

However, we found no evidence for the hypothesis that system-paced learning would aggravate the negative effect of presenting unnecessary text on learning, as we found no differences among layout conditions when learning was self- or system paced.

In Experiment 2, which was system-paced, conducted in a lab, and involved eye tracking to study participants' processing of the unnecessary information, we did replicate the finding that a separated presentation (without integration instruction) led to better learning outcomes than an integrated presentation of unnecessary text. However, none of the unnecessary text conditions showed lower learning outcomes than the diagram-only condition. Moreover, although participants in the integrated condition spent about as much time looking at the unnecessary text as participants in the separated and integration instruction conditions, they did make more transitions between the unnecessary text and the relevant parts of the diagram. A possible, but speculative, explanation is that the unnecessary text was not harmful for learning when it is merely read, without deep processing or integration attempts, but will negatively influence learning when learners attempt to actively integrate the unnecessary information with the relevant information.

One potential explanation for the failure to replicate Chandler and Sweller's (1991) findings on the effects of unnecessary text on learning might lie in their relatively small sample size (i.e., between ten and fourteen participants per condition). Indeed, given the presence of publication bias, the positive predictive value of a significant result (i.e., the probability that a significant effect represents a true effect) decreases with the sample size. Yet, their study was not the only one showing that extraneous information can hamper learning (for reviews: see Mayer & Fiorella, 2014, on the coherence principle; and Kalyuga & Sweller, 2014, on the redundancy principle). However, what qualified as 'extraneous information' differed among studies. Because both the relation between the extraneous information and the essential information (i.e., whether the extraneous information is irrelevant or unnecessary), and the modality in which the extraneous information is presented might influence the occurrence and strength of the negative effect on learning, further studies should be performed taking into account these factors. That is, it is possible that the negative effects of irrelevant information (i.e., not related to the learning goal; e.g., Harp & Mayer, 1998; Moreno & Mayer, 2000; Sanchez & Wiley 2006) are larger than the effects of unnecessary information (i.e., related to the learning goal, but not necessary because the information is presented twice or is unnecessarily elaborate; Bobis, Sweller, & Cooper, 1993; Chandler & Sweller, 1991; Mayer et al., 1996). Moreover, the strength of the effect may depend on whether it is text (Bobis et al., 1993; Chander & Sweller, 1991; Mayer et al., 1996), picture (Sanchez & Wiley, 2006) or sound that is extraneous (Moreno & Mayer, 2000).

Such studies should also employ eye tracking to study the attention allocation processes during learning. Indeed, theory predicts that extraneous information hampers learning because learners select this information, attend to it, and try to integrate it with the essential information. This presumably hampers learning by depleting valuable working memory resources which cannot be used for processes relevant for learning (Kalyuga & Sweller, 2014, Sweller et al., 2011). The present study indeed suggests that the layout of the learning materials also influences the integration process. While participants in all unnecessary-text conditions attended to the text, participants in the integrated layout tried to integrate the unnecessary text with the relevant information more.

Note though, that if such integration attempts negatively affect learning, one would also expect performance to be lower compared to a diagram only condition, which was not the case. This might be associated with a potential limitation of Experiment 2: the fact that it was system-paced with a 100s presentation time (i.e., 20s longer than in Experiment 1), may have meant that some participants had more time available than they needed or would otherwise have taken to study the learning materials (e.g., the self-paced separated condition only took 64s on average in Experiment 1), whereas others may not have had sufficient time (e.g., the self-paced diagram-only condition took 129s on average in Experiment 1, though with large variability between participants).

In sum, the present study shows no evidence for a negative effect of unnecessary information on learning compared to a diagram-only condition. However, a separated presentation led to better learning outcomes than an integrated presentation of unnecessary text, and participants made more transitions between the unnecessary text and the relevant parts of the diagram when it was integrated. This suggests that when text is presented that describes a picture, it is better not to integrate it into the picture, but to present it separatel

Chapter 6

Summary and Discussion

Research focusing on the design of multimedia materials has shown that the presentation of extraneous information hampers learning. The research presented in this dissertation focussed on the question of whether this negative effect of extraneous information on learning with multimedia materials would decrease or disappear with task experience (i.e., familiarity with the design of the materials). Extraneous information can be either irrelevant or unnecessary for attaining the learning goals. Irrelevant information is not related to the learning goal (e.g., 'bells and whistles' added to learning materials to make them interesting or motivating; known as seductive details, see Mayer, Heiser, & Lonn, 2001). Unnecessary information in contrast, is related to the learning goal, but not necessary for learning, for instance because the information is presented twice in a different form (e.g. text and picture, Chandler & Sweller, 1991; spoken and written text, Kalyuga, Chandler, & Sweller, 1999) or is unnecessarily elaborate (e.g., details and examples that are not necessary for the learning goals; Reder & Anderson, 1982).

A large body of research has demonstrated that the presentation of extraneous information, whether irrelevant or unnecessary, hampers learning from multimedia materials (for reviews, see Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). However, eye-tracking literature has shown that people may learn to ignore task-irrelevant information with increasing expertise (Canham & Hegarty, 2010; Haider & Frensch, 1999). Therefore, the studies presented in this dissertation examined whether the negative effect of extraneous information on learning would decrease or disappear with increasing task experience. When extraneous information is present over a series of tasks, items, or slides, learners have a chance to adapt their study strategy (if they realize the information is irrelevant or unnecessary). When learners would start to ignore the extraneous information, its negative effect on learning should diminish or disappear.

Therefore, the aim of this dissertation was to provide an answer to the following questions: 1) Does the negative effect of extraneous information on learning decrease or disappear with increasing task experience? 2) Does this effect arise because learners start to ignore the extraneous information? These two questions were addressed both for irrelevant (Part 1) and unnecessary (Part 2) information presentation. In this chapter, the key findings of the studies are summarized and discussed in the context of theories of cognitive load and multimedia learning, and potential directions for future research are sketched.

Summary of Key Findings

The studies in **Part 1**, presented in Chapters 2 and 3, investigated whether the negative effect of *irrelevant* information on learning would decrease or disappear with increasing task experience. **Chapter 2** described three experiments on the effects of irrelevant pictures when learning action-word definitions. In all experiments, participants learned the definitions of new words (from an artificial language) that denoted actions, coupled with matching pictures (depicting the same action), mismatching pictures (depicting another action), or without pictures. Experiments 1a/1b addressed the question whether adding irrelevant information (the mismatching pictures) would hamper learning of word definitions compared to learning words with matching pictures or without pictures. Results showed lower test performance (i.e., definition recall) when learning with mismatching pictures than with matching pictures. Experiment 2 examined the main hypothesis that irrelevant information would no longer hamper learning once learners gained experience with the word-learning task. As expected, learning with the mismatching pictures led to significantly lower test performance compared to learning with the matching pictures initially (replicating findings from Experiment 1). However, after participants gained experience with the task, test performance in the mismatching condition improved and no longer differed significantly from the matching pictures condition. Experiment 3 employed eye tracking to test the hypothesis that the negative effect on learning disappeared with increasing task experience in Experiment 2, because learners started to ignore the irrelevant information. In line with this hypothesis, eye-tracking measures showed that the increase in recall performance in the mismatching pictures condition was accompanied by a decrease in attention towards these pictures. This suggests that participants spontaneously (i.e., without any instruction to do so) adapted their study strategy, and stopped allocating attention towards the mismatching pictures.

The study presented in **Chapter 3** built on the findings from Chapter 2, by investigating whether learners 'only' ignored the location of the mismatching pictures or had learned to suppress attention to the content of the mismatching pictures. This question was addressed in Experiments 1a/1b, by systematically changing the location of matching and mismatching pictures for half of the participants after they had accumulated task experience. If participants would only suppress attention to the picture *location*, then word learning should be negatively affected after the location change for participants engaging in word learning with mismatching pictures.

However, if they were aware that the *content* of the mismatching pictures was irrelevant for learning the word definitions, they would be expected to actively suppress attention to the pictures regardless of the location, in which case performance should not be negatively affected after the location change for participants engaging in word learning with mismatching pictures. Results showed that changing the location of the mismatching pictures did not affect recall, suggesting awareness of the irrelevance of the pictures for their learning goal.

Concluding, the studies presented in Part 1 showed that learners seemed to be able to adapt their study strategy after they gained task experience: they started to ignore the irrelevant information (at which point their learning outcomes improved) and seem to do so because they are aware that it is irrelevant for their learning goal.

The studies in **Part 2**, presented in Chapters 4 and 5, investigated the negative effect of *unnecessary* information on learning, using more complex (i.e., higher element interactivity) multimedia materials (i.e., expository text with pictures on biological processes).

In the two experiments in **Chapter 4**, the hypothesis was addressed that learners would start to ignore unnecessary textual information (which merely described the picture) with increasing task experience, thereby reducing its negative effects on learning. In addition, it was investigated whether the layout of the unnecessary information (integrated in or separated from the picture) would influence the effect of unnecessary text on attention and learning over time (i.e., with task experience). Participants learned about the process of mitosis with materials consisting of a combination of essential text and pictures (control), essential text and pictures with unnecessary text presented either integrated in or separated from the picture. It was hypothesized that an initial negative effect of unnecessary information would occur; that this negative effect would decrease (or even disappear) as participants gained task experience; and that this decrease would be stronger when the unnecessary text was presented separated from the picture (i.e., separated unnecessary text would be easier to ignore than integrated unnecessary text). Surprisingly, the presentation of unnecessary text did not consistently hamper learning about the process of mitosis in the two experiments. Because we did not reliably find an initial negative effect of the unnecessary text on learning, the question of whether task experience would reduce or eliminate that negative effect could not be answered. However, the eye-movement data collected in Experiment 2 did show that the unnecessary textual information was processed by students, and more interestingly, that they seemed to start ignoring this unnecessary information with increasing task experience.

In the two experiments described in **Chapter 5** participants studied multimedia materials about the functioning of the heart. This study involved a replication of Experiment 5 of Chandler and Sweller (1991), which was part of a series of experiments showing that (1) the addition of unnecessary text to a self-containing diagram impeded learning; (2) that this negative effect of unnecessary text on learning was larger when the text was integrated in the diagram than when the text was presented separately; and (3) that this negative effect on learning was larger when participants were instructed to mentally integrate the separated unnecessary information with the diagram as compared to no such instruction. A second aim of the first experiment in Chapter 5 was to examine the influence of pacing (system-paced vs. self-paced) on the occurrence and size of the negative effect of unnecessary information, as this negative effect might be larger when learning is system-paced, because under system-paced conditions, any unnecessary information processing goes directly at the expense of time available for essential information processing.

In these experiments, we used four different lay-outs in four different conditions: 1) a diagram presented without unnecessary text (diagram only); 2) a diagram with unnecessary text separated from the diagram (separated); 3) a diagram with unnecessary text separated from the diagram with the instruction to mentally integrate the text and the diagram (integration instruction); or 4) a diagram with unnecessary text integrated into the diagram (integrated). In Experiment 1 study time was self-paced for half of the participants, and system-paced for the other half.

We hypothesized that the unnecessary text would have a negative effect on learning when it was integrated in the essential material, or when participants received the integration instruction. We also addressed the open question whether participants would be able to ignore the spatially separated unnecessary text to such an extent that it would not hamper their learning compared to a diagram-only condition. In addition, we hypothesized that system-pacing would aggravate the negative effects of unnecessary text on learning.

Over the two experiments, the presentation of unnecessary text did not hamper learning about the functioning of the heart compared to a diagram-only condition, regardless of whether participants learned at their own pace or under system-paced conditions. In Experiment 2, we did replicate the finding from Chandler and Sweller (1991) that a separated presentation (without integration instruction) led to better learning outcomes than an integrated presentation of unnecessary text. The eye-tracking data indicated that this was not because participants in the integrated condition spent more time on the unnecessary information: they spent about as much time looking at the unnecessary text as participants in the separated and integration instruction conditions but they presumably attempted to integrate the unnecessary text and the essential parts of the diagram more (as evidenced by more transitions). This might explain why in the integrated condition, the unnecessary text interfered more with learning than in the separated condition without integration instruction.

Discussion of Key Findings

The results of the studies presented in Part 1 of this dissertation extend prior multimedia research by directly testing the prevailing explanations for the negative effect of irrelevant information on learning. That is, our findings show that such information initially captures attention (i.e. is being processed). This presumably hinders learning by drawing on valuable working memory resources that can no longer be used for processes relevant for learning (Kalyuga & Sweller, 2014; Sweller et al., 2011) and/or by actively conflicting with the to be learned information (cf. Mayer et al., 2007). Using eye-tracking methodology, we showed that a reduction in attention to irrelevant information was associated with an increase in recall performance, providing some evidence that irrelevant information indeed captures attention and working memory resources during learning. This assumption, that irrelevant information captures attention, is rarely directly testing in studies on multimedia learning, and direct tests of assumptions about the (attentional) mechanisms underlying multimedia principles are important, because they can support existing ideas about how and why effects on learning occur, or may generate new insights and explanations for these principles (Van Gog & Scheiter, 2010).

These findings extend prior eye-tracking research, which already showed that people learn to ignore task-irrelevant information during task *performance* as a result of task experience (Haider & Frensch, 1999) or explicit instruction (Canham & Hegarty, 2010; Hegarty et al., 2010). Part 1 of this dissertation showed that people can also learn to ignore irrelevant information during *learning*, and that this is associated with better learning outcomes. That the negative effect of extraneous information diminishes when people gain task experience is interesting and seemingly contradicts the expertise reversal effect (Kalyuga, 2014). The expertise reversal effect states that learning materials that are essential and non-redundant for novices, become redundant when learners gain or have more prior knowledge, at which point they will no longer aid, and might even hinder learning. Thus, an expertise reversal effect would imply that information becomes extraneous and *starts* to hamper learning as task experience increases, whereas in our studies the extraneous information (in this case irrelevant) *stops* hampering learning as task experience increases. Moreover, in contrast to the expertise reversal effect, participants in our studies did not gain experience with (or knowledge of) the *task content* (i.e., the word definitions to be learned) over time (as there were no repetitions, new words were being presented throughout the experiments), but with the task *presentation*.

That being said, it does seem to be the case that students learn to ignore irrelevant information because they are aware that this information does not help their learning, and not solely based on the location of the screen at which it is presented. This implies that either 1) top-down attentional influences (i.e., conscious redirection of attention because a stimulus does not contain relevant information) are stronger than bottom-up attentional influences (i.e., saliency of a stimulus suddenly appearing at a different location; cf. Remington, Johnston, & Yantis, 1992), or 2) that initial bottom-up attentional influences are suppressed quickly (i.e., the *signal-suppression hypothesis*; cf. Gaspelin, Leonard, & Luck, 2015; Sawaki & Luck, 2010). Either way, it seems that people are capable of a conscious adaptation of their attention allocation, and start to attend more to the relevant information that is presented on the screen, irrespective of the location of the irrelevant information. This shows that task experience might be a boundary condition to the negative effects of irrelevant information as learners seem to be able to adapt their study strategy and start to ignore irrelevant information.

While the studies of Part 1 show the expected pattern of results for extraneous information that was irrelevant and pictorial in nature, the same does not hold true for the studies in Part 2 with extraneous information that was unnecessary and textual. While the accumulation of task experience did lead to changes in the allocation of attention (i.e., participants attended less to the unnecessary information with increasing task experience in Chapter 4), this did not lead to the expected improvement in learning outcomes. Because we found no consistent evidence that the unnecessary information had an initial negative effect on learning in both Chapters reported in Part 2, it was impossible to study the effect of task experience on the negative effect of unnecessary information on learning. This failure to find a consistent negative effect of unnecessary information is surprising, considering the body of research suggesting that this is a quite robust effect (for reviews, see Kalyuga &

Sweller, 2014; Mayer & Fiorella, 2014) over a wide variety of instructional material (i.e., identical information in written and spoken form, unnecessary text accompanying diagram, or unnecessary details and examples). The negative effect of unnecessary text on learning has also been shown in various prior studies (e.g., Bobis, Chandler, & Sweller, 1993; Chandler & Sweller, 1991; Pociask & Morrison, 2008), yet we were unable to replicate these findings.

A potential explanation might lie in the amount of time that learners had available to process and learn the materials in both studies. In Chapter 4, we based the time per slide on the average study time of eight participants in a pilot study, while in Chapter 5, we based the study time for the participants in the system-paced condition on the average study time of the participants in the Chandler and Sweller (1991) study. These study times can therefore be assumed to have been sufficient for most of our participants, an assertion that is supported by the finding that learning in a system-paced environment did not significantly hamper learning compared to a self-paced environment in Chapter 5. It is possible that processing unnecessary information would start to hamper learning when there is time pressure. When there is little time available for processing, any time spent on the unnecessary information goes at the expense of thoroughly processing essential information, and as a result, learning is hampered. In these studies, learners may have had sufficient time for processing all sources of information, which would explain why their attention to the unnecessary text did not significantly increase experienced cognitive load and did not negatively affect learning.

Despite the differences between irrelevant and unnecessary information (to which we will return in the next section), the studies presented in this dissertation are relevant in the context of relatively recent lines of research on cognitive load and multimedia learning. Whereas research has long been focused on what multimedia designers can do to aid student learning, the present studies focused on whether students can (learn) to adapt their study strategy and ignore sources of extraneous cognitive load. This fits well with a recent line of research in which participants are instructed to self-manage their cognitive load during learning by changing their study strategy (e.g., Agostinho, Tindall-Ford, & Roodenrys, 2013; Gordon, Tindall-Ford, Agostinho, & Paas, 2016; Roodenrys, Agostinho, Roodenrys, & Chander, 2012), and with research on training multimedia learning strategies (e.g., Bodemer, Ploetzner, Feuerlein, & Spada, 2004; Mason, Pluchino, & Tornatora, 2016; Peshkam, Mensink, Putnam, & Rapp, 2011; Stalbovs, Scheiter, & Gerjets, 2015). For instance Roodenrys et al. (2012) showed that students were able to self-manage their cognitive load under conditions of split attention by physically manipulating the instructional materials to alleviate the need for integration (the split-attention effect occurs when a student is required to divide his attention between two or more mutually referring sources; Ayres & Sweller, 2014). Moreover, Peshkam et al. (2011) showed that negative effects of unnecessary text that was added to an expository text were alleviated to some extend by warning readers that such information might be present in the text they were about to read. In contrast to those studies, in which participants were actively instructed or trained to manage their own cognitive load, the results from Chapters 2 and 3 (performance and eye tracking data) and Chapter 4 (eye tracking data) indicate that students are also able to implicitly adapt their study strategy and improve their learning outcomes.

Our findings that students are able to ignore extraneous information with increasing task experience during learning are also interesting in the context of recent studies regarding selective attention (e.g., Middlebrooks, Kerr, & Castel, 2017), which can compensate for negative effects of divided attention during learning (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). In the study of Middlebrooks et al. (2017) participants had to learn words while either performing an auditory dual task (i.e., responding to different types of tones), or listening to music. Participants were awarded more points for some words than for others, making the words with higher rewards more valuable. Results showed that participants could partly overcome a negative effect of dividing one's attention between the two tasks by selectively attending to the most valuable words. This shows that students can also re-allocate their attention based on the value of the information they have to learn. In the studies presented in this dissertation, participants implicitly adapted their study strategy, so they decided themselves which information held more value for attaining their learning goal.

These findings on self-management of cognitive load and instructed or spontaneous adaptation of study strategies are also of interest for educational practice, for instance for instructional designers of multimedia learning materials. That is, it is very hard for instructional designers to take into account all multimedia principles, because individual learner characteristics may interact with some of the principles. For example, information that essential for novices, might be redundant for more experience learners (i.e., the expertise-reversal effect, Kalyuga, 2014). Some examples of expertise-reversal include explanations with diagrams might be crucial for novices, while these become redundant as they gain expertise (Kalyuga, Chandler, & Sweller, 2000), and whereas novices benefit from studying worked examples, more advanced learners are better off solving practice problems (Kalyuga, Chandler, & Sweller, 2001). However, when students are able to self-manage their cognitive load, instructional designers do not have to develop tailor-made instructional materials for students with different backgrounds and expertise.

Should the finding that learners are able to ignore irrelevant information with increasing task experience be replicated in a wider range of learning situations, then instructional materials could be enriched with irrelevant but interesting information (i.e. seductive details) to enhance students' interest and motivation. The negative effects of irrelevant information on learning might be smaller than expected in the long run, as students might only briefly glance at this information after they gain experience with the instructional materials, rendering the negative effects of such information on learning negligible, while it might still have a motivating effect, in which case such information might even improve learning in the long run when it successfully increases situational interest (cf. Magner, Schwonke, Aleven, Popescu, & Renkl, 2014), or enjoyment (cf. Lenzner, Schnotz, & Müller, 2013). However, it should be noted that it is unsure that our findings would translate to other types of irrelevant information, as the pictures used in Part 1 were not only irrelevant for learning the action word definition, but might also have actively interfered with learning, by depicting another action. An interesting question is whether it would be more difficult for participants to learn to ignore irrelevant information that does not actively interfere with the to-be-learned information, as the lack of conflict might make it less obvious that the information can/should be ignored.

Limitations and Future Research

The results of this dissertation provide some evidence that irrelevant information more consistently harms learning as compared to unnecessary information. These differences might find their origin in the relation between the extraneous information and the learning goal, which is profoundly different for unnecessary and irrelevant information. Indeed, irrelevant information is not related to the learning goal, and as such is extraneous for all learners irrespective of the expertise or prior knowledge the learner possesses. Unnecessary information, however, is related to the learning goals, but is unnecessary because a better representation of the information is available or one source of information is already self-explaining. It is possible that the negative effects of irrelevant information on learning would be larger than the effects of unnecessary information. That is, processing irrelevant information not only takes up working memory capacity, but might actively interfere with learning the essential information, by disrupting the processing, organization, and integration of essential information. Processing unnecessary information on the other hand (which is identical in content to the essential information), does take up working memory capacity but may interfere less with learning the essential information –unless the available study time is limited. Thus, it would be interesting for future research to also address the role of time pressure: Unnecessary information might mostly hamper learning when available study time is limited, while irrelevant information can also hamper learning when it is self-paced.

However, multiple other possible reasons exist as to why different patterns of results emerged over the two parts of this dissertation. Firstly, the studies presented in the two parts differed in the complexity of the tasks that participants had to complete, as they had to learn words in Part 1 (which is lower in task complexity, with fewer interacting elements), while they had to learn about biological processes in Part 2 (which is higher is task complexity, with more interacting elements). However, based on CLT we would sooner expect to find negative effects of extraneous information in task with a higher complexity, as these tasks would already pose a high intrinsic cognitive load, and, thus, an increase in extraneous load would lead to a working memory overload. It should be noted though, that the word learning task used in Part 1, although low in complexity (with few interacting elements), was cognitively demanding for our participants, as shown by the amount of effort invested (high), and the performance on the cued recall test (low).

Another difference between the two parts of this dissertation lies in the manner in which the irrelevant and unnecessary information is presented. While the irrelevant information in Part 1 was pictorial, the unnecessary information in Part 2 was textual. Because there are differences in attention allocation to text and pictures (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Hannus & Hyönä, 1999; Schmidt-Weigand, Kohnert, & Glowalla, 2010), it would be interesting to investigate whether the medium in which irrelevant or unnecessary information is presented, would affect attention and learning.

Finally, the results of this dissertation show that task experience might be a boundary condition to the negative effects of irrelevant information on learning as students seem to be able to adapt their study strategy and start to ignore irrelevant information. However, other possible boundary conditions seem to exist, as this negative effect seems to depend on individual difference in attentional control (Rey, 2014) and working memory (Sanchez & Wiley, 2006). To this date, these individual differences are rarely controlled for in studies regarding design of multimedia instructional materials, including the studies presented in this dissertation. These concepts are related to core assumptions of both CLT (e.g., a cognitive overload is less likely when a student has more working memory resources available) and the CTML (e.g., a student with more attention control is less likely to select irrelevant information over essential information). Therefore, future research on the design of multimedia learning material in general and on extraneous information processing in particular, should aim to include measures of attention control and working memory. The identification of boundary conditions is important for scientific progress as they the limits of generalizability of scientific theories (Busse, Kach, & Wagner, 2016; Whetten, 1989). In other words, when boundary conditions are uncovered, they can explain previously unexpected results, and open up new avenues for research. Hopefully, the results of the studies presented in this dissertation will do just that.

References

Agostinho, S., Tindall-Ford, S., & Roodenrys, K. (2013). Adaptive diagrams: Handing control over to the learner to manage split-attention online. *Computers and Education, 64*, 52-62. doi:10.1016/j.compedu.2013.01.007

Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *American Journal of Psychology, 117*, 389-410. doi:10.2307/4149007

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 940-945. doi:10.1037/a0029199

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, rev. ed.) (pp.206-226). New York: Cambridge University Press. doi:/10.1017/CBO9781139547369.011

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423. doi:10.1016/S1364-6613(00)01538-2

Barouillet, P., & Camos, V. (2007). The time-based resource-sharing model of working memory. In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 59-80). Oxford: Oxford University Press

Bauhoff, V., Huff, M., & Schwan, S. (2012). Distance matters: Spatial contiguity effects as trade-off between gaze-switches and memory load. *Applied Cognitive Psychology, 26*, 863-871. doi:10.1002/acp.2887

Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction, 3*, 1-21. doi:10.1016/S0959-4752(09)80002-9

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

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology, 40*, 109-131. doi:10.1146/annurev.ps.40.020189.000545

Buhrmester, M., Kwang, T., & Gosling, S.D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3-5. doi:10.1177/1745691610393980

Busse, C., Kach, A. P., & Wagner, S. M. (2016). Boundary conditions: What they are, how to explore them, why we need them, and when to consider them. *Organizational Research Methods*. doi:10.2139/ssrn.2713980

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, rev. ed.) (pp.174-206). New York: Cambridge University Press. doi:10.1017/CBO9781139547369.010

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*, 155-166. doi:10.1016/j.learninstruc.2009.02.014

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General, 138*, 351-367. doi:10.1037/a0015854

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332. doi:10.1207/s1532690xci0804_2

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory and Cognition, 29*, 1146-1152. doi:10.3758/BF03206384

Chi, M. T. H., DeLeeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477. doi:10.1207/s15516709cog1803_3

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.), New Jersey: Lawrence Erlbaum Associates

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-114. Doi:10.1017/S0140525X01003922

Craig, S., Gholson, B., & Driscoll, D. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*, 428-434. doi:10.1037/0022-0663.94.2.428

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. C. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125*, 159-180. doi:10.1037/0096-3445.125.2.159

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagram. *Contemporary Educational Psychology, 35*, 59-74. doi:10.1016/j.cedpsych.2009.10.002

De Nooijer, J. A., Van Gog, T., Paas, F., & Zwaan, R. A. (2013). When left is not right: Handedness effects on learning object-manipulation words using pictures with left- or right-handed first-person perspectives. *Psychological Science, 24*, 1-7. doi:10.1177/0956797613498908

Farley, A. P., Ramonda, K., & Liu, X. (2014). The concreteness effect and the bilingual lexicon: The impact of visual stimuli attachment on meaning recall of abstract L2 words. *Language Teaching Research, 16*, 449-466. doi:10.1177/1362168812436910

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychological Science, 26*, 1740-1750. doi:10.1177/0956797615597913

Gordon, C., Tindall-Ford, S., Agostinho, S., & Paas, F. (2016). Learning from instructor-managed and self-managed split-attention materials. *Applied Cognitive Psychology, 30*, 1-9. doi:10.1002/acp.3159

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology, 30*, 304-337. doi:10.1006/cogp.1996.0009

Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 172-190. doi:10.1037/0278-7393.25.1.172

Hald, L. A., Van Den Hurk, L., & Bekkering, H. (2015). Learning verbs more effectively through meaning congruent action animations. *Learning and Instruction, 39*, 107-122. doi:10.1016/j.learninstruc.2015.05.010

Hannus, M. & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24*, 95-123.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*, 414-434. doi:10.1037/0022-0663.90.3.414

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301-307. doi:10.1016/S0896-6273(03)00838-9

Hegarty, M., Canham, M., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 37-53. doi:10.1037/a0017683

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphic: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology, 23*, 1215-1226. doi:10.1002/acp.1525

Jamet, E., & Le Bohec, O. (2007). The effect of redundant text in multimedia instruction. *Contemporary Educational Psychology, 32*, 588-598. doi:10.1016/j.cedpsych.2006.07.001

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. doi:10.1016/j.learninstruc.2009.02.019

Kahneman, D. (1973). *Attention and effort*. New Jersey: Prentice-Hall.

Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, rev. ed.) (pp.576-597). New York: Cambridge University Press. doi:10.1017/CBO9781139547369.028

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371. doi:10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Education Psychology, 92*, 126-136. doi:10.1037//0022-0663.92.1.126

Kalyuga, S., Chandler, P., & Sweller, J. (2001). Learner experience and efficiency of instructional guidance. *Educational Psychology, 21*, 5-23. doi:10.1080/01443410124681

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46*, 567-581. doi:10.1518/hfes.46.3.567.1640

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd, rev. ed.) (pp.247-262). New York: Cambridge University Press. doi:10.1017/CBO9781139547369.10

Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174. doi:10.2307/2529310

Lehman, S., Schraw, G., McCrudden, M., & Hartly, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology, 32*, 568-587. doi:10.1016/j.cedpsych.2006.07.002

Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science, 41*, 811-831. doi:10.1007/s11251-012-9256-z

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Technology Research and Development, 30*, 195-232. doi:10.1007/BF02765184

Little, J. L., & Bjork, E. L. (2016). Multiple-choice pretesting potentiates learning of related information. *Memory and Cognition, 44*, 1085-1101. doi:10.3758/s13421-016-0621-z

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education, 5*, 196-211. doi:10.1111/j.1751-228X.2011.01129.x

Magner, U. I. E., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction, 29*, 141-152. doi:10.1016/j.learninstruc.2012.07.002.

Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. British Journal of Educational Technology, 47,1083-1095. doi: 10.1111/bjet.12271

Mayer, R. E. (Ed.) (2014). *The Cambridge handbook of multimedia learning* (2nd, rev. ed.). New York: Cambridge University Press. doi:10.1017/CBO9781139547369

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*, 64-73. doi:10.1037/0022-0663.88.1.64

Mayer, R. E., DeLeeuw, K. E., & Ayres, P. (2007). Creating retroactive and proactive interference in multimedia learning. *Applied Cognitive Psychology, 21*, 795-809. doi:10.1002/acp.1350

Mayer, R. E., & Fiorella, L. (2014). 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* (2nd, rev. ed.) (pp.279-315). New York: Cambridge University Press. doi:10.1017/CBO9781139547369.015

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in in less understanding. *Journal of Educational Psychology, 93*, 187-198. doi:10.1037/0022-0663.93.1.187

Mayer, R. E., & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding. *Journal of Experimental Psychology: Applied, 11*, 13-18. doi:0.1037/1076-898X.11.1.13

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology, 100*, 380-386. doi:10.1037/0022-0663.100.2.380

Mayer, R., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43-52. doi:10.1207/S15326985EP3801_6

Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively distracted: Divided attention and memory for important information. *Psychological Science, 28*, 1103-115. doi:10.1177/0956797617702502

Miller, G. (1956). The magic number seven, plus or minus two: Some limits to our capacity for processing information. *Psychological Review, 63*, 81-97.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92*, 117-125. doi:10.1037//0022-0663.92.1.117

Nasiopoulos, E., Risko, E. F., Foulsham, T., & Kingstone, A. (2015). Wearable computing: Will it make people prosocial? *British Journal of Psychology, 106*, 209-216. doi:10.1111/bjop.12080

Oksa, A., Kalyuga, S., Chandler, P. (2010). Expertise reversal effect in using explanatory notes for readers of Shakespearean text. *Instructional Science, 38*, 217-236. doi:10.1007/s11251-009-9109-6

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. doi:10.1037/0022-0663.84.4.429

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1-4. doi:10.1207/S15326985EP3801_1

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71. doi:10.1207/S15326985EP3801_8

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411-419.

Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect: A multimedia study. *Computers in Human Behavior, 27*, 5-10. doi:10.1016/j.chb.2010.05.006

Peshkam, A., Mensink, M. C., Putnam, A. L., & Rapp, D. N. (2011). Warning readers to avoid irrelevant information: When being vague might be valuable. *Contemporary Educational Psychology, 36*, 219-231. doi:10.1016/j.cedpsych.2010.10.006

Pociask, F. D., & Morrison, G. R. (2008). Controlling split attention and redundancy in physical therapy instruction. *Educational Technology Research and Development, 56*, 379-399. doi:10.1007/s11423-007-9062-5

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience, 21*, 793-797. doi:10.1111/j.1460-9568.2005.03900.x

Reder, L. M., & Anderson, J. R. (1982). Effects of spacing and embellishment on memory for the main points of a text. *Memory and Cognition, 10*, 97-102. doi:10.3758/BF03209210

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics, 52*, 279-290. doi:10.3758/BF03212254

Rey, G. D. (2014). Seductive details and attention distraction – An eye tracker experiment. *Computer in Human Behavior, 32*, 133-144. doi:10.1016/j.chb.2013.11.017

Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: Implied social presence, eye tracking and attention. *Attention, Perception and Psychophysics, 73*, 291-296. doi:10.3758/s13414-010-0042-1

Roodenrys, K., Agostinho, S., Roodenrys, S., & Chandler, P. (2012). Managing one's own cognitive load when evidence of split attention is present. *Applied Cognitive Psychology, 26*, 878-886. doi:10.1002/acp.2889

Rop, G., Van Wermeskerken, M., De Nooijer, J. A., Verkoeijen, P. P. J. L., & Van Gog, T. (in press). Task experience as a boundary condition to the negative effects of irrelevant information on learning. *Educational Psychology Review.* Advance online publication. doi:10.1007/s10648-016-9388-9

Rop, G., Verkoeijen, P. P. J. L., & Van Gog, T. (2017). With task experience students learn to ignore the content, not just the location of irrelevant information. *Journal of Cognitive Psychology, 29*, 599-606. doi:10.1080/20445911.2017.1299154

Sadoski, M. (2005). A dual coding view of vocabulary learning. *Reading & Writing Quarterly, 21*, 221-238. doi:10.1080/10573560590949359

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory and Cognition, 34*, 344-355. doi:10.3758/BF03193412

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics, 72*, 1455-1470. doi:10.3758/APP.72.6.1455

Scheiter, K., Schüler, A., Gerjets, P., Huk, T., & Hesse, F. W. (2014). Extending multimedia research: How do prerequisite knowledge and reading comprehension affect learning from text and pictures. *Computers in Human Behavior, 31*, 73-84. doi:10.1016/j.chb.2013.09.022

Schmidt-Weigand, F., Kohnert, A. & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20*, 100-110. doi: 10.1016/j.learninstruc.2009.02.011

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

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior, 29*, 1590-1601. doi:10.1016/j.chb.2013.01.047

Shen, H. H. (2010). Imagery and verbal coding approaches in Chinese vocabulary instruction. *Language Teaching Research, 14*, 485-499. doi:10.1177/1362168810375370

Smith, B. D., Stahl, N., & Neil, J. (1987). The effect of imagery instruction on vocabulary development. *Journal of College Reading and Learning, 20*, 131-137. doi:10.1080/10790195.1987.10849909

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

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312. doi:10.1016/0959-4752(94)90003-5

Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review, 22*, 123-138. doi:10.1007/s10648-010-9128-5

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer: New York. doi:10.1007/978-1-4419-8126-4

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17*, 273-281. doi:10.1162/0898929053124965

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*, 16-26. doi:10.1080/00461520701756248

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. doi:10.1002/acp.1112

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20,* 95-99. doi:10.1016/j.learninstruc.2009.02.009

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, *14*, 490-495. Doi:10.5465/AMR.1989.4308371

Wilkins, D. (1972). *Linguistics in language teaching*. London: Arnold.

Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty? *Journal of Educational Psychology, 105*, 266-277. doi:10.1037/a00319

Nederlandse samenvatting

Summary in Dutch

Onderzoek heeft aangetoond dat het presenteren van overbodige informatie in multimediaal lesmateriaal[1] een negatief effect heeft op de leerprestatie. De studies gepresenteerd in dit proefschrift, richtten zich op de vraag of dit negatieve effect van overbodige informatie op leren van multimediaal lesmateriaal zou verminderen of verdwijnen naarmate lerenden ervaring met het materiaal opbouwen (oftewel, bekend raken met het design van deze materialen). Er zijn twee subtypen van overbodige informatie te onderscheiden; deze informatie kan ofwel *irrelevant* ofwel *onnodig* zijn voor het bereiken van de leerdoelen. Irrelevante informatie is niet gerelateerd aan het leerdoel (bijv. ‘toeters en bellen’ die worden toegevoegd aan instructiematerialen om het ‘op te leuken’, met als doel de interesse en motivatie van de lerende te verhogen; zie Mayer, Heiser, & Lonn, 2001). Onnodige informatie is wel gerelateerd aan het leerdoel, maar niet noodzakelijk voor het leerproces, bijvoorbeeld omdat dezelfde informatie twee keer wordt gepresenteerd in verschillende vormen (bijv. als tekst en als afbeelding, Chandler & Sweller, 1991; zowel in gesproken als geschreven tekst, Kalyuga, Chandler, & Sweller, 1999), of omdat het onnodig uitgebreid is (bijv. details en voorbeelden in een tekst die niet nodig zijn om de kern te begrijpen; Reder & Anderson, 1982).

Vele wetenschappelijke studies hebben aangetoond dat het toevoegen van overbodige informatie, of het nou irrelevante of onnodige informatie is, het leren met multimediaal lesmateriaal schaadt (voor reviews, zie Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014). Echter, ‘eye-tracking’ onderzoek, waarin de oogbewegingen van participanten werden gemeten, heeft laten zien dat mensen kunnen leren om taak-irrelevante informatie te negeren naarmate ze meer ervaring met een taak opdoen (Canham & Hegarty, 2010; Haider & Frensch, 1999). Gebaseerd op deze schijnbare tegenstelling, stond in de studies gepresenteerd in dit proefschrift de vraag centraal of het negatieve effect van overbodige informatie op de leerprestatie zou afnemen of verdwijnen naarmate lerenden ervaring met het materiaal opbouwen. Mogelijk zijn lerenden in staat om hun leerstrategie aan te passen wanneer overbodige informatie telkens aanwezig is tijdens het werken aan taken, of het bestuderen van items of slides. Wanneer lerenden na enige tijd de overbodige informatie zouden gaan negeren, zou

[1] Multimediaal lesmateriaal is lesmateriaal bestaande uit een combinatie van tekst en beeld, waarbij de tekst zowel geschreven als gesproken kan zijn en het beeld zowel statisch (bijvoorbeeld foto’s of grafieken) als dynamisch (bijvoorbeeld video of animatie) kan zijn

het negatieve effect dat deze informatie op hun leerprestatie heeft, moeten afnemen of verdwijnen.

Het doel van dit proefschrift was dan ook om een antwoord te geven op de volgende vragen: 1) Vermindert of verdwijnt het negatieve effect van overbodige informatie op de leerprestatie met toenemende ervaring met het lesmateriaal? 2) Treedt dit effect op doordat lerenden de overbodige informatie gaan negeren? In dit proefschrift is een poging gedaan om deze twee vragen te beantwoorden voor zowel het presenteren van irrelevante (Deel 1) en onnodige (Deel 2) informatie in multimediaal lesmateriaal.

De studies in **Deel 1**, gepresenteerd in Hoofdstuk 2 en 3, onderzochten of het negatieve effect van *irrelevante* informatie op de leerprestatie zou verminderen of verdwijnen naarmate de ervaring van de lerenden met het lesmateriaal toeneemt. In **Hoofdstuk 2** worden drie experimenten beschreven over de effecten van irrelevante afbeeldingen bij het leren van definities van actiewoorden. In alle experimenten leerden de participanten de definities van nieuwe woorden (uit een kunstmatige taal) die een handeling beschrijven. De geschreven woorden en gesproken definities werden vergezeld door afbeeldingen die de actie uitbeeldden (hierna matchende afbeeldingen genoemd), afbeeldingen die een andere actie uitbeeldden (hierna mismatchende afbeeldingen genoemd), of werden zonder afbeelding gepresenteerd. In Experimenten 1a/1b werd onderzocht of het toevoegen van irrelevante informatie (d.w.z. de mismatchende afbeeldingen) het leren van de woorddefinities zou schaden ten opzichte van het leren van dezelfde definities met matchende afbeeldingen of zonder afbeeldingen. De resultaten lieten zien dat participanten lagere testscores behaalden wanneer zij leerden met de mismatchende afbeeldingen dan wanneer zij leerden met de matchende afbeeldingen. In Experiment 2 werd de belangrijkste hypothese onderzocht, namelijk dat irrelevante informatie het leren niet langer zou schaden wanneer participanten ervaring hadden opgedaan met de woordleertaak. Zoals verwacht, leidde leren met de mismatchende afbeeldingen aanvankelijk tot significant lagere testscores in vergelijking met leren met matchende afbeeldingen (deze bevinding repliceert die van Experiment 1). Echter, nadat de participanten ervaring hadden opgedaan met de woordleertaak, nam de prestatie in de conditie met mismatchende afbeeldingen toe en was er geen verschil meer in testscores tussen de condities met mismatchende en matchende afbeeldingen. In Experiment 3 werden de oogbewegingen van participanten gemeten om de hypothese te testen dat het negatieve effect van irrelevante informatie op leren verdween in Experiment 2 doordat lerenden de mismatchende afbeeldingen begonnen te negeren. De resultaten lieten het verwachte beeld zien: De toename van testscores in de conditie met mismatchende afbeeldingen ging gepaard met een afname in aandacht voor deze afbeeldingen. Dit suggereert dat participanten spontaan (oftewel zonder enige instructie) hun studiestrategie aanpasten en geen aandacht meer schonken aan de mismatchende afbeeldingen.

De studies gepresenteerd in **Hoofdstuk 3,** bouwden voort op de bevindingen uit Hoofdstuk 2, door te onderzoeken of lerenden enkel de locatie van de mismatchende afbeeldingen begonnen te negeren, of dat ze hun aandacht onderdrukten op basis van de inhoud van de deze afbeeldingen. In Experimenten 1a/1b werd voor de helft van de participanten de locatie van de matchende en de mismatchende afbeeldingen systematisch veranderd nadat ze ervaring met de woordleertaak hadden opgedaan. Als participanten enkel hun aandacht voor de de mismatchende afbeeldingen zouden onderdrukken op basis van de *locatie*, dan zou de leerprestatie van participanten die leerden met mismatchende afbeeldingen negatief beïnvloed moeten worden door de locatieverandering. Echter, wanneer het voor de participanten duidelijk was dat de *inhoud* van de afbeeldingen irrelevant was voor het leren van de woorddefinities, dan zouden zij in staat moeten zijn om hun aandacht voor deze afbeeldingen te onderdrukken ongedacht de locatie van de afbeeldingen en zou hun leerprestatie niet negatief beïnvloed moeten worden door de locatiewijziging. Uit de resultaten bleek dat de wijziging van de locatie van de mismatchende afbeeldingen de testscores niet beïnvloedde, wat suggereert dat participanten zich ervan bewust waren dat deze afbeeldingen irrelevant waren voor hun leerdoel.

Kortom, de studies uit Deel 1 laten zien dat lerenden in staat zijn om op basis van hun ervaring met het lesmateriaal hun leerstrategie aan te passen, waarbij ze irrelevante informatie gaan negeren waardoor hun leerprestaties toenemen. Het lijkt erop dat zij die informatie gaan negeren omdat zij zich ervan bewust zijn dat deze irrelevant is voor hun leerdoel.

De studies in **Deel 2**, gepresenteerd in Hoofdstuk 4 en 5, onderzochten het effect van *onnodige* informatie op de leerprestatie. In deze studies werd gebruikt gemaakt van meer complex multimediaal lesmateriaal (d.w.z. een verklarende tekst met relevante afbeeldingen over biologische processen, waarin meerdere informatie elementen met elkaar in verband gebracht moeten worden om van de tekst te leren).

In de twee experimenten in **Hoofdstuk 4** werd de hypothese onderzocht dat lerenden onnodige tekstuele informatie (die enkel de afbeelding beschreef) zouden gaan negeren nadat ze meer ervaring hadden opgedaan met het materiaal, waarna een initieel negatief effect op leren zou verminderen of verdwijnen. Daarnaast werd onderzocht of de lay-out van de onnodige informatie (d.w.z. geïntegreerd in of gescheiden van de afbeelding) invloed zou hebben op het effect van de onnodige tekst op de aandacht en leerprestatie met toenemende ervaring. Participanten leerden over het mitoseproces met materialen die bestonden uit een combinatie van essentiële tekst en afbeeldingen (controle conditie), of essentiële tekst en afbeeldingen met onnodige tekst, die ofwel geïntegreerd was in de afbeelding ofwel los van de afbeelding gepresenteerd werd. De verwachting was dat een initieel negatief effect van onnodige tekst op de leerprestatie zou optreden; dat dit negatieve effect zou verminderen (of zelfs verdwijnen) wanneer participanten ervaring met het lesmateriaal op zouden doen; en dat deze afname sterker zou zijn wanneer de onnodige tekst los van de afbeelding gepresenteerd werd (omdat deze makkelijker te negeren zou moeten zijn dan geïntegreerde tekst). Opvallend genoeg vonden we geen consistent negatief effect van de toegevoegde onnodige tekst op het leren over het mitoseproces in de beide experimenten. Doordat er geen initieel negatief effect op de leerprestatie optrad, kon ook de vraag of dit effect zou verminderen of verdwijnen met toenemende ervaring met het lesmateriaal niet beantwoord worden. Echter, de oogbewegingsdata uit Experiment 2 lieten wel zien dat lerenden aandacht schonken aan de onnodige tekst en dat ze –overeenkomstig onze verwachting- minder aandacht aan deze onnodige tekst gingen besteden naarmate hun ervaring met het materiaal toenam.

In de twee experimenten beschreven in **Hoofdstuk 5,** werd gebruik gemaakt van multimediaal lesmateriaal over het functioneren van het hart. Deze studie was gedeeltelijk een replicatie van Experiment 5 van Chandler en Sweller (1991), die in een serie experimenten lieten zien dat (1) de toevoeging van onnodige tekst aan een afbeelding van het hart (of een elektrische schakeling) die ook zonder de tekst te begrijpen was, een negatief effect had op de leerprestatie; (2) dat dit negatieve effect van de onnodige tekst op de leerprestatie groter was wanneer de tekst geïntegreerd was in de afbeelding dan wanneer de tekst los daarvan gepresenteerd werd; en (3) dat dit negatieve effect op de leerprestatie groter was wanneer participanten geïnstrueerd werden om los van de afbeelding gepresenteerde onnodige tekst mentaal te integreren met de afbeelding, dan wanneer zij geen integratie-instructie kregen. Een tweede doel van het eerste experiment in Hoofdstuk 5 was om de invloed van studietijd (d.w.z. vaststaande studietijd of in eigen tempo) op het optreden en de grootte van het negatieve effect van onnodige informatie op de leerprestatie te onderzoeken. De verwachting was dat dit negatieve effect groter zou zijn bij een vaststaande studietijd, omdat het verwerken van onnodige informatie dan direct ten koste gaat van de tijd die beschikbaar is om de essentiële informatie te verwerken.

In deze experimenten werd gebruik gemaakt van verschillende lay-outs in de vier condities: 1) een afbeelding van het hart zonder onnodige tekst (enkel afbeelding), 2) een afbeelding met onnodige tekst die los van de afbeelding gepresenteerd werd (gescheiden), 3) een afbeelding met onnodige tekst die los van de afbeelding gepresenteerd werd met de instructie om de tekst en de afbeelding mentaal te integreren (integratie instructie), of 4) een afbeelding met onnodige tekst die geïntegreerd in de afbeelding gepresenteerd werd (geïntegreerd). In Experiment 1 leerde de helft van de participanten in elke conditie op hun eigen tempo, terwijl de andere helft leerde met een vaststaande studietijd. De hypothese was dat er een negatief effect van de onnodige tekst op de leerprestatie zou optreden in de condities waarin deze tekst geïntegreerd was in de afbeelding of wanneer lerenden de integratie instructie kregen vergeleken met de controle (enkel afbeelding) conditie. Daarnaast onderzochten we de open vraag of participanten in de 'gescheiden tekst zonder integratie instructie' conditie in staat zouden zijn om de gescheiden tekst te negeren, waardoor deze de leerprestatie niet zou schaden in vergelijking met het leren van enkel de afbeelding. Ten slotte verwachtten we dat leren met een vaststaande studietijd het negatieve effect van onnodige tekst op de leerprestatie zou versterken.

In beide experimenten leidde het toevoegen van onnodige tekst niet tot slechtere leerprestaties dan leren met enkel de afbeelding. Het maakte hierbij niet uit of de participanten leerden in hun eigen tempo, of met een vaststaande studietijd. In Experiment 2 repliceerden wij wel de bevinding van Chandler en Sweller (1991) dat leren met los van de afbeelding gepresenteerde onnodige tekst (zonder integratie instructie) leidde tot betere leerprestaties dan leren met in de afbeelding geïntegreerde onnodige tekst. De oogbewegingsdata lieten zien dat dit niet kwam doordat participanten in de geïntegreerde conditie meer tijd besteedden aan de onnodige tekst: deze participanten besteedden eenzelfde hoeveelheid tijd aan de onnodige tekst als participanten in de gescheiden tekst en de gescheiden tekst met integratie instructie condities. Echter, participanten in de geïntegreerde tekst conditie leken meer pogingen te doen om de onnodige tekst mentaal te integreren met de essentiële delen van de afbeelding (blijkens meer transities tussen tekst en afbeelding), wat mogelijk kan verklaren waarom de onnodige tekst in de geïntegreerde conditie meer interfereerde met het leren dan in de gescheiden conditie zonder de integratie instructie.

In **Hoofdstuk 6** worden de resultaten bediscussieerd. Uit de resultaten van dit proefschrift kunnen twee belangrijke conclusies getrokken worden, namelijk: (1) met toenemende ervaring met het lesmateriaal zijn lerenden in staat om overbodige informatie te negeren, en zich meer te focussen op de essentiële informatie en (2) of dit ook effect heeft op de leerprestaties lijkt te verschillen tussen irrelevante en onnodige overbodige informatie. Een mogelijke verklaring voor het verschil in effect op de leerprestaties tussen irrelevante en onnodige informatie ligt in de relatie van deze informatie met het leerdoel. Aangezien irrelevante informatie niet gerelateerd is aan het leerdoel, kan het verwerken van deze informatie actief interfereren met het leren van de essentiële informatie. Onnodige informatie, daarentegen, is gerelateerd aan het leerdoel, en interfereert daarom mogelijk niet direct met het leren van de relevante informatie. Echter hebben eerdere studies wel een negatief effect van onnodige informatie op leerprestaties gevonden. Vermoedelijk zal onnodige informatie het leren wel gaan schaden wanneer de beschikbare studietijd zeer beperkt is. Hierin ligt ook een mogelijk limitatie van Deel 2 van dit proefschrift, aangezien de participanten (zelfs in de condities met vaststaande studietijd) waarschijnlijk ruim voldoende tijd hadden om alle informatie te verwerken, waardoor we geen negatief effect van onnodige informatie op leren konden vaststellen.

Met betrekking tot de eerste conclusie, tevens de voornaamste bevinding van dit proefschrift, lijkt het erop dat ervaring met het lesmateriaal een randvoorwaarde kan zijn voor het wel of niet optreden van het negatieve effect van overbodige informatie op leren (in elk geval voor wat betreft irrelevante overbodige informatie). Dit omdat mensen overbodige informatie leren te negeren, waardoor het negatieve effect op de leerprestatie vermindert of zelfs verdwijnt. Het zoeken naar zulke randvoorwaarden waaronder effecten wel of niet optreden is belangrijk voor de wetenschap, omdat men zo de grenzen van de generaliseerbaarheid van wetenschappelijke theorieën en bevindingen kan vaststellen (Busse, Kach, & Wagner, 2016; Whetten, 1989).

Tevens is deze bevinding mogelijk van belang voor de onderwijspraktijk –hoewel verder onderzoek wenselijk en noodzakelijk is. Aan veel instructiematerialen wordt namelijk irrelevante informatie toegevoegd (de eerder genoemde ‘toeters en bellen’, ofwel ‘verleidelijke details’; Harp & Mayer, 1998) om de materialen aantrekkelijker en meer motiverend te maken voor lerenden (Lenzner, Schnotz, & Müller, 2013; Magner, Schwonke, Aleven, Popescu, & Renkl, 2014). Omdat uit veel eerder onderzoek bleek dat zulke informatie het leren schaadt (voor reviews, zie Kalyuga & Sweller, 2014; Mayer & Fiorella, 2014), werd aanbevolen dat instructiematerialen niet verrijkt moeten worden met zulke ‘verleidelijke details’. Deze eerdere onderzoeken hielden echter weinig rekening met de ontwikkeling van aandachts- en cognitieve processen over de tijd naarmate ervaring met het materiaal toeneemt. Wanneer lerenden in staat zouden blijken te zijn om bij een breed scala van instructiematerialen de irrelevante informatie grotendeels te negeren met toenemende ervaring, zal het negatieve effect op leren van zulke informatie mogelijk kleiner zijn dan verwacht. Als lerenden, na het opbouwen van taakervaring, maar kort naar de aantrekkelijke details kijken, kan het toevoegen van zulke informatie aan instructiematerialen het leren mogelijk zelfs verbeteren door de interesse en het leerplezier van lerenden te verhogen.

Dankwoord (Acknowledgements)

Ik wist niet of ik hier ooit aan toe zou komen, maar nu de rest van mijn proefschrift af is moet ik er toch aan geloven; een dankwoord. Met een aan zekerheid grenzende waarschijnlijkheid durf ik te stellen dat dit het meest gelezen deel van mijn proefschrift zal worden, al is het maar omdat iedereen opzoekt of, en hoe, hij of zij bedankt wordt. Aangezien het in mijn ogen ondoenlijk is om iedereen in voldoende mate te bedanken, ben ik dan ook huiverig naar de reacties op dit dankwoord. Dit gezegde hebben ga ik toch een poging wagen.

Ten eerste wil ik Tamara en Peter bedanken voor hun begeleiding tijdens dit promotietraject. Tamara, in mijn ogen was jij een dagelijks begeleider en promotor in een. Zonder jouw ondersteuning, aanmoedigingen, en kritische blik op mijn werk was dit proefschrift nooit in deze vorm tot stand was gekomen. Peter, jouw inhoudelijk iets andere, en daardoor verfrissende, kijk op de theorieën gebruikt in dit proefschrift, en je uitgebreide kennis van de statistiek, hebben dit proefschrift naar een hoger niveau geholpen. Ook wil ik jullie beide ontzettend bedanken dat ik qua intensiteit van begeleiding weinig hebt gemerkt van jullie verhuizingen naar respectievelijk Utrecht (Tamara) en Breda (Peter).

Het werk als promovendus was nooit zo leuk geweest wanneer ik niet het geluk had gehad om terecht te komen in een groep geweldige collega's. Daarom wil ik graag de volledige Human Learning and Performance groep bedanken voor de gezelligheid en de sfeer van verbondenheid. Ook wil ik graag de leden van de pubgroep bedanken, voor de constructieve feedback en stimulerende gesprekken, en voor de gezelligheid en het plezier dat daarna volgde in de Smitse. Daarnaast wil ik de hele groep psychologie AIO's bedanken, voor de vele phd-diners, borrelmomenten, en zaalvoetbalavonturen. In het bijzonder wil ik hier graag Charly, Gerdien, Jesper, Marien, Mario, Marit, Steven, Tim, Vincent, en Wim bedanken.

Ook wil ik graag mijn kamergenoten bedanken voor de geweldige tijd die we samen hebben gehad. Margina en Keri, bedankt voor het warme welkom op de kamer, het delen van de teleurstelling van een afgewezen stuk, maar ook de blijdschap van een publicatie. Dat dit laatste gepaard ging met een kindersurprise ei maakt de overwinning extra zoet. Sander, ook jou wil ik bedanken voor je gezelligheid en je hulp bij allerhande zaken.

Een (aspirerend) wetenschapper kan niet zonder ondersteuning, zowel tijdens de dagelijkse gang van zaken als bij het uitvoeren van het onderzoeken. Daarom wil

ik graag iedereen van zowel het secretariaat als het Erasmus Behavioral Lab bedanken voor hun hulp en ondersteuning.

Voor mijn verblijf in Tübingen wil ik graag iedereen in AG6 bedanken voor de leuke en productieve tijd die ik daar heb gehad. Voornamelijk Katharina Scheiter en Anne Schüler wil ik bedanken voor enerzijds hun gastvrijheid en anderzijds de begeleiding tijdens en na mijn verblijf.

Ook wil ik deze ruimte graag gebruiken om Rob Withagen te bedanken. Rob was de persoon die mijn sluimerende interesse voor de wetenschap wist aan te spreken, en zonder zijn begeleiding tijdens mijn bachelor- en masterproject had ik nooit gekozen voor een academische carrière.

Ik was nooit opgegroeid tot de persoon die ik nu ben zonder de steun en liefde van mijn ouders en broers en zus. Als jongste kind had ik de bevoorrechte positie dat alle paden al voor mij gebaand waren, en dat ik nooit iets écht fout kon doen. Pap, Mam, Marja, en Niels, bedankt!

Ook heb ik het geluk gehad om vele vriendschappen opgebouwd te hebben tijdens mijn leven. Zowel in Stadskanaal, Groningen, als Rotterdam heb ik altijd enorm genoten van alle feestjes, borrels, barbecues, spelletjesavonden, etc. Hierbij wil ik bij uitstek Chrisjon en Ernst bedanken voor hun vriendschap. Ik denk dat de band die wij hebben erg bijzonder is, en die wil ik graag koesteren in de toekomst.

Het laatste dankwoord wil ik graag reserveren voor de belangrijkste persoon in mijn leven: Martine. Dankzij jou weet ik hoe het voelt om onvoorwaardelijke liefde te geven en te ontvangen. Daarnaast hebben jouw steun, motivatie, en tips ontzettend geholpen tijdens het afronden van mijn promotie. Het feit dat ik mijn verdere leven met jou mag delen maakt van mij de gelukkigste mens op aarde.

Curriculum Vitae

Gertjan Rop was born on December 15, 1987 in Stadskanaal. In 2006 he obtained his Athenaeum degree from secondary school Ubbo Emmius lyceum in Stadskanaal, and started his bachelor Human Movement Sciences at the University of Groningen. Because of his growing interest in the grounding of differences between experts and novices in sport, his bachelor thesis examined the influence of neuropsychological capacities on level and type of sport. Thereafter, he started the Sports, Learning and Performance master, and during two years was trained in all facets of being a researcher. He graduated in 2012, and his thesis, which focused on the development of expertise, and why some individuals become experts at a task while others do not, was published in *Attention, Perception, & Psychophysics* (Rop & Withagen, 2014). In 2014, he started his three year PhD research project in Educational Psychology at the Department of Psychology, Education, and Child Studies at the Erasmus University Rotterdam.

In his PhD project that resulted in this dissertation, Gertjan studied how people learn with multimedia learning materials. More specifically, his research revolved around the redundancy and coherence principles, which state that inclusion of extraneous information hampers learning with multimedia materials. He studied experimentally, using learning outcomes and eye-tracking methodology, whether people can learn to ignore such information with increasing experience with the task. During his PhD trajectory, Gertjan was a visiting researcher in Prof. Dr. Katharina Scheiter's group at the Leibniz-Institut für Wissensmedien (IWM) in Tübingen (Germany). Currently, Gertjan is employed as a postdoctoral researcher in the Educational Psychology group at the Erasmus University Rotterdam. He is continuing and expanding his research into learning with multimedia materials, now focussing on the question of how the split-attention effect is best explained. Next to his PhD and postdoc research activities, Gertjan has been tutoring in courses on statistics, multimedia learning, and instructional design, and (co)supervising bachelor and master theses students

Publications

Published

Rop, G., Verkoeijen, P. P. J. L., & Van Gog, T. (2017). With task experience students learn to ignore the content, not just the location of irrelevant information. *Journal of Cognitive Psychology, 29*, 599-606 doi:10.1080/20445911.2017.1299154

Rop, G., Van Wermeskerken, M., De Nooijer, J. A., Verkoeijen, P. P. J. L., & Van Gog, T. (2016). Task experience as a boundary condition to the negative effects of irrelevant information on learning. *Educational Psychology Review*. Advance online publication. doi: 10.1007/s10648-016-9388-9

Rop, G., & Withagen, R. (2014). Perceivers vary in their capacity to benefit from feedback in learning to perceive length by dynamic touch. *Attention, Perception, & Psychophysics, 76,* 864-876. doi:10.3758/s13414-013-0598-7

Submitted

Pouw, W. T. J. L., **Rop, G.**, & Paas, F. (2017). *The cognitive basis for the split-attention effect*. Manuscript submitted for publication.

Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2017). *Effects of task experience and layout on learning from text and pictures with or without unnecessary picture descriptions.* Manuscript submitted for publication.

Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2017). *The effect of layout and pacing on learning from diagrams with unnecessary text*. Manuscript submitted for publication.

Presentations

Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (September, 2017). *Effects of task experience and layout on learning from text and pictures with or without unnecessary picture descriptions*. Paper presented at EARLI 2017, the 17th biennial Conference of the European Association for Research on Learning and Instruction, Tampere, Finland.

Rop, G. (November, 2017). *Negative effects of extraneous information on learning with multimedia*. Invited presentation, Leibniz-Institut für Wissensmedien, Tübingen, Germany.

Rop, G., Verkoeijen, P. P. J. L., & Van Gog, T. (July, 2016). *Negative effects of irrelevant information on learning disappear: People learn to ignore the content, not the location.* Poster presented at Special Interest Group 2 (Comprehension of Text and Graphics) meeting of EARLI, Geneva, Switzerland.

Rop, G., Verkoeijen, P. P. J. L., & Van Gog, T. (June, 2016). *Negative effects of irrelevant information on learning disappear because people learn to ignore the content, not just the location*. Paper presented at the 9th annual Cognitive Load Theory Conference, Bochum, Germany.

Rop, G. (Februari, 2016). *The redundancy effect and task experience.* Invited presentation, Leibniz-Institut für Wissensmedien, Tübingen, Germany.

Rop, G., Van Wermeskerken, M., De Nooijer, J. A., Verkoeijen, P. P. J. L., & Van Gog, T. (September, 2015). *The redundancy effect disappears with task experience.* Poster presented at the Embodied Cognition in Multimedia Learning symposium, Rotterdam, the Netherlands.

Rop, G., De Nooijer, J. A., Verkoeijen, P. P. J. L., & Van Gog, T. (August, 2015). *The redundancy effect in artificial language word learning.* Poster presented at EARLI 2015, the 16th biennial Conference of the European Association for Research on Learning and Instruction, Limassol, Cyprus.

Rop, G., Van Wermeskerken, M., Verkoeijen, P. P. J. L., & Van Gog, T. (August, 2015). *The redundancy effect in word learning: Students learn to ignore irrelevant pictures.* Paper presented at the JURE 2015, the pre-conference of the Junior Researchers at EARLI 2015, the 16th biennial Conference of the European Association for Research on Learning and Instruction, Limassol, Cyprus.