Eur J Psychol Educ (2014) 29:483–501
DOI 10.1007/s10212-013-0209-y

Focus of attention and choice of text modality in multimedia learning

Wolfgang Schnotz · Christoph Mengelkamp · Christiane Baadte · Georg Hauck

Received: 18 May 2013 / Revised: 8 December 2013 / Accepted: 11 December 2013 /
Published online: 15 January 2014

Abstract The term 'modality effect' in multimedia learning means that students learn better from pictures combined with spoken rather than written text. The most prominent explanations refer to the split attention between visual text reading and picture observation which could affect transfer of information into working memory, maintenance of information in working memory or the effective size of working memory. The assumption of a continuous need for split attention is questionable, however. Learners can keep pictorial information in working memory, when they have seen the picture before, especially if they have higher prior knowledge. Instead of suffering from a permanent split attention, learners frequently show tendencies to simply ignore pictures. This suggests guiding learners towards picture analysis by picture-related text paragraphs. We assume that these paragraphs are associated with stronger modality effects than content-related paragraphs, especially if the pictures are new to learners. These assumptions were tested in an experiment with 120 students learning about volcanism from illustrated text consisting of segments each including a content-related paragraph followed by a picture-related paragraph describing the accompanying visualization. Content-related and picture-related paragraphs were presented as visual or auditory texts leading to 2x2 conditions of text presentation. Picture novelty was manipulated by presenting a picture throughout the whole segment or only when the picture-related paragraph was read. As expected, picture-related paragraphs were associated with stronger modality effects than

W. Schnotz (✉) · C. Baadte
Faculty of Psychology, General and Educational Psychology, University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany
e-mail: schnotz@uni-landau.de

C. Baadte
e-mail: baadte@uni-landau.de

C. Mengelkamp
Philosophical Faculty II, University of Würzburg, Oswald-Külpe-Weg 82, 97074 Würzburg, Germany
e-mail: christoph.mengelkamp@uni-wuerzburg.de

G. Hauck
Communication and Environment, Rhine-Waal University of Applied Sciences, Südstraße 8, 47475 Kamp-Lintfort, Germany
e-mail: georg.hauck@hochschule-rhein-waal.de

content-related paragraphs if picture novelty is high. The distinction between different kinds of paragraphs seems to be important for the prediction of modality effects.

Keywords Modality effect · Multimedia learning · Split-attention · Content-related paragraphs · Picture-related paragraphs

Modern multimedia learning environments allow flexible combinations of visualizations with written or spoken language. Much research has been invested into the question of how verbal information should be presented along with pictorial information. Various studies have found that students learn better from pictures combined with spoken rather than written text (e.g. Jeung et al. 1997; Low and Sweller 2005; Mayer 2009; Mayer and Moreno 1998; Moreno and Mayer 1999; Mousavi et al. 1995; Schmidt-Weigand et al. 2010a, b; Tindall-Ford et al. 1997; cf. Ginns 2005, for a meta-analysis). These findings have been referred to as modality effects.

Beyond multimedia learning, there is a long tradition of research focusing on general differences between auditory and visual text unrelated to pictures (Danks 1980; Massaro 1979; Rubin et al. 2000; Sticht and James 1984). Auditory texts include prosody, but have a transient nature (Danks and End 1987). Visual texts lack prosody, but are relatively stable. They allow learners to adapt their reading pace to the requirements of cognitive processing which is advantageous especially for difficult texts (Ferreira and Anes 1994; Green 1981; Hron et al. 1985). For short texts, listening seems to result in better learning than reading (cf. Rummer et al. 2010). For longer texts, listeners seem to better focus on the main points (i.e. macro-propositions), whereas readers seem to concentrate more on the details (i.e. micro-propositions) (Kürschner and Schnotz 2008; Rubin et al. 2000). Diakidoy et al. (2012) have recently emphasized that differences between learning from reading and learning from listening decrease with increasing reading skills.

Modality effects in multimedia learning can be found only under specific conditions. They seem more likely to occur if the learning content is complex or if the presentation is system-paced rather than self-paced or if learners' prior knowledge is low (Gyselinck et al. 2008; Leahy and Sweller 2011; Mayrath et al. 2011; Stiller et al. 2009; Tindall-Ford et al. 1997). In various studies, novices performed better after learning from a visual display plus an auditory text rather than a visual display plus visual text. However, the effect reversed when learners had higher expertise (Kalyuga 2000;Kalyuga et al. 2000; Mousavi et al. 1995; Schmidt-Weigand 2011; Tabbers and van der Spoel 2011). Leahy and Sweller (2011) found a modality effect only when text length was short assuming that this differential effect was due to the transient nature of auditory text. All in all, the modality effect in multimedia learning is not based on a unitary cognitive mechanism. It aggregates superficially similar effects that result from different causes (Schnotz 2011; Schüler et al. 2011).

The following paper argues that there is no reason for assuming that students learn generally better from pictures combined with spoken rather than written text. Young as well as adult learners frequently ignore pictures while reading text as they underestimate the pictures' informational content (e.g. Hannus and Hyönä 1999; Mokros and Tinker 1987; McDonald and Thornley 2002; Weidenmann 1989). Thus, various authors have recommended special guidance for learning from text and pictures to enhance learners' picture analysis (Kozma 2003; Lowe 1999, Leahy et al. 2003; Moreno and Durán 2004). Such guidance can be provided through extra text paragraphs aiming at a deeper picture analysis by telling learners where to look and what to see (Ploetzner et al. 2013). These additional picture-related text paragraphs should require more split attention and, thus, be associated with stronger modality effects than the regular content-related paragraphs. After an analysis of possible moderating

factors influencing modality effects, the following paper will report an experiment aimed at testing this assumption.

Theory

Split attention

The most straightforward assumption for why pictures with spoken text are better for learning than pictures with written text refers to the avoidance of split attention. The learner's eyes cannot simultaneously focus on two different sources of information. In order to use both sources, the reader of an illustrated text has to split his/her visual attention by switching his/her eyes back and forth between words and pictures which is not needed with auditory text and pictures. Avoiding split attention matters if multiple external information sources are really needed for learning. If learners have sufficient prior knowledge, single sources of information can also be understood in isolation. Chandler and Sweller (1991) demonstrated that adding a text to a diagram which is intelligible by itself does not improve learning. Kalyuga et al. (1998) found that when learners' expertise increases, physically integrating multiple sources of information (as a means to minimize split attention) first lost its advantage and then became disadvantageous in comparison to a physically separated presentation (Kalyuga 2000; Kalyuga et al. 2003, 2000; Mousavi et al. 1995; Yeung et al. 1997). There are two different reasons for a modality effect due to split attention, one referring to the transmission of information from the eyes and ears through *sensory channels to working memory*, and the other referring to the simultaneous availability of *verbal and pictorial information within working memory* as a requirement for integrative cognitive processing (Schnotz 2005; Schnotz, in press).

Information transfer into working memory Both the visual sensory channel and the auditory sensory channel have limited transmission capacity. Only a limited amount of information can be transferred into working memory within a specific time interval. Because two sensory channels provide more capacity, the overall information input into working memory per time unit will be higher with two channels than with only one channel. If the visual channel is used for pictorial information and the auditory channel is used for verbal information and if learning time is limited to an amount that is just sufficient to convey the relevant information through the two channels, it follows that the usage of only one (i.e. the visual) channel under the same time constraints makes it impossible to transfer all the relevant information into working memory: The learner will necessarily miss some information which will not be available for mental model construction (Schüler et al. 2011). It should be noted that this aspect of split attention is simply a lack of information due to an impeded information transfer to working memory because of too severe time constraints. Such a lack of information is less likely to occur, when the visual channel is devoted to the picture and the auditory channel is devoted to the text. To summarize, a modality effect can be produced by limiting the available learning time due to the fact that the learner will necessarily miss some information when operating with only one sensory channel under too severe time constraints.

Information availability in working memory Another explanation of modality effects by split attention focuses on the simultaneous availability of verbal *and* pictorial information in working memory. Working memory is limited both in terms of (a) capacity and (b) storage time: (a) only a limited amount of information can be simultaneously

held in working memory; (b) information fades from working memory within 20 seconds if not maintained by continuous processing (cf. Cowan 2001; Miller 1956; Peterson and Peterson 1959). Multimedia learning can therefore be considered as the construction of mental models from a transitory verbal and pictorial data base in working memory.

Coherence formation between verbal and pictorial information requires that both kinds of information are simultaneously available in working memory. Due to the high decay-rates of working memory, split attention necessarily reduces simultaneous availability of verbal and pictorial information in working memory (Ayres and Sweller 2005). Simultaneous availability can be enhanced by spatial or temporal contiguity of information. Spatial contiguity means that written text should be presented in close spatial proximity to the picture because this reduces visual search processes when shifting attention from the text to the picture and vice versa. Temporal contiguity means that spoken text should be presented simultaneously with the picture (Mayer 2005, 2009).

Both types of contiguity serve the same purpose: to increase simultaneous availability of verbal and pictorial information in working memory. Spatial contiguity can never totally avoid split attention because the eyes still have to fluctuate between the text and the picture. Split attention can only be fully avoided by temporal contiguity using the auditory and the visual modality. In this case, the learner can devote his/her full visual attention to the picture and his/her full auditory attention to the text. Thus, avoidance of split attention by presenting pictures with spoken text means that a maximum of verbal and pictorial information can be simultaneously held available in working memory.

It should be remembered that according to numerous experiments, students learn better from text and pictures than from text alone because texts and pictures allow cross-referential connections between both kinds of information and constructing more elaborate mental representations (Levie and Lentz 1982; Levin et al. 1987). This has been called the multimedia effect (Mayer 1997, 2005, 2009). Against this background, the rationale of the modality effect is maximizing the multimedia effect. Under this condition, using the auditory and the visual modality implies maximizing the temporal contiguity of verbal and pictorial information in working memory. It also implies minimizing the negative effects of split attention by reducing the obstacles for simultaneous availability of information in working memory. Thus, a modality effect is only to be expected if there is also a multimedia effect. If there is no need for integrating verbal and pictorial information, no modality effect is to be expected either (Jeung et al. 1997; Leahy et al. 2003; Schnotz 2011).

Size of working memory

Whereas the split attention explanation of modality effects deals with information transfer and with simultaneous availability of information in working memory, another explanation postulates variations in the effective size of working memory. Moreno and Mayer (1999) presented text and pictures to learners in a consecutive way which avoided split attention. Nevertheless, spoken text with pictures resulted in better learning than written text with pictures. The authors argued that effective working memory capacity varies with modalities. When written text and pictures are processed, only visual working memory is involved. When spoken text and pictures are processed, both visual and auditory working memory is involved. According to this view, verbal and pictorial information can be integrated more effectively when more memory capacity is involved, even if the two systems receive their input in a consecutive manner.

Although plausible at first sight, this view is problematic on two accounts. First, not only auditory but also visual text is processed in auditory memory, because readers recode graphemic text information into phonological information (Baddeley 1986; Baddeley and Logie 1999; Ellis and Young 1996; Rieben and Perfetti 1991). Second, the text material used by Mayer and Moreno consisted mostly of single sentences. Rummer et al. (2010) have argued that the effects found by Moreno and Mayer result from the longer duration of auditory sensory traces, which plays an important role when only sentences are presented, rather than the usage of more working memory capacity (Lindow et al. 2011).

Limiting conditions

Prior knowledge Students do not learn better from text and pictures under all conditions. Prior knowledge can totally eliminate the multimedia effect. If learners have sufficient prior knowledge, the learning content can be intelligible for them also without pictures. Therefore, a combination of text and pictures which has positive effects for learners with low prior knowledge can have negative effects on learning when prior knowledge is high (Kalyuga et al. 2000). If a picture is not needed for comprehension, split attention is unlikely to occur (Leahy et al. 2003, experiment 2) and, thus, no modality effect is to be expected either.

Picture familiarity Even if comprehension benefits from pictures as an external support for mental model construction, it does not mean that learners have to continuously look at the picture. If a picture has a very simple structure that can be easily perceived and remembered, the relevant information can be held in working memory also without continuous external support. Similarly, if the learner has seen a picture previously and the picture is sufficiently simple, there is no need for repeatedly re-visiting the picture as external support for mental model maintenance in working memory. Instead, the learner can read the text without split attention and nevertheless interconnect verbal and pictorial information based on his/her visual image or mental model. In short, there is no reason to assume a need for a continuous pictorial support and no reason for an ongoing need for split attention. A modality effect is only to be expected when learners need the pictures to a considerable extent for mental model construction. Low prior knowledge learners who receive complex pictures not seen before can be expected to need the pictures to a considerable extent, because these pictures need repeated visual investigation to convey the relevant information into working memory. On the contrary, learners with high prior knowledge who receive simple pictures or ones they are already familiar with will not have much use for these pictures.

Non-consideration of pictures Split attention explanations of modality effects are only valid to the extent that learners are willing to use both the text and the corresponding picture. Previous research on learning from written text and pictures by children and by adults has shown that this is often not the case (Hannus and Hyönä 1999; Mokros and Tinker 1987; McDonald and Thornley 2002; Weidenmann 1989). Many students underestimate the informational value of pictures. They process visual displays only superficially even if there are no time constraints. These students suppose that a passing view of a picture is sufficient to grasp the essential information. From this perspective, there is not so much split attention but a wide-spread non-consideration of the pictures that is to be avoided. Presenting pictures with auditory text offers learners the possibility of listening to the text while simultaneously investigating the picture. With a visual text, the picture is likely to receive much less attention even if there is sufficient learning time available.

Research question and hypotheses

Because students often underestimate the informational value of pictures, researchers have recommended guidance for learners in order to enhance their picture analysis (Bernard 1990; Kozma 2003; Kulhavy et al. 1985; Lowe 1999, 2004; Ploetzner et al. 2013). Learners can receive guidance through extra text paragraphs (beyond the ordinary text) that accompany the pictures. These paragraphs explain how the picture represents its content, give directions of how to analyze the picture and make the results of the analysis explicit. Generally speaking, these extra paragraphs guide learners towards a deeper picture analysis by telling them where to look and what to see (Hegarty and Just 1993). If text authors follow this recommendation, a distinction can be made between two kinds of text paragraphs: content-related paragraphs and picture-related paragraphs. Content-related paragraphs deal directly with the learning content. For example, if a text deals with plate tectonics, the following is a content-related paragraph:

(1) Plate tectonics is a scientific theory that describes slow large-scale motions in the Earth crust usually measured in centimeters per year. Among these motions, one type - called 'subduction' - frequenly leads to the formation of volcanoes. Subduction means that one tectonic plate moves under another tectonic plate. As the plates converge, the lower plate sinks into the Earth mantle, where it partially melts due to high temperatures. The resulting magma rises towards the surface, where it forms volcanoes …

This text paragraph is all about the learning content: about the movement of tectonic plates, about subduction and its consequences. On the contrary, picture-related paragraphs do not directly deal with the learning content. They deal with the picture and how the content is visualized. The following is an example of a picture-related paragraph (referring to the picture in Fig. 1) dealing with how the content is visualized in the picture:

(2) Consider the picture shown in Fig. 1. It shows as a cross-section through the earth's mantle which processes occur when one tectonic plate, visualized by the left dark band (Oceanic Crust), slides under another tectonic plate, visualized by the right thick dark band (Continental Crust). The bubbles above the dark band sliding downwards (Oceanic Crust) show areas of partial melting turning into magma. The upwards arrows indicate the rise of the magma towards the earth surface …

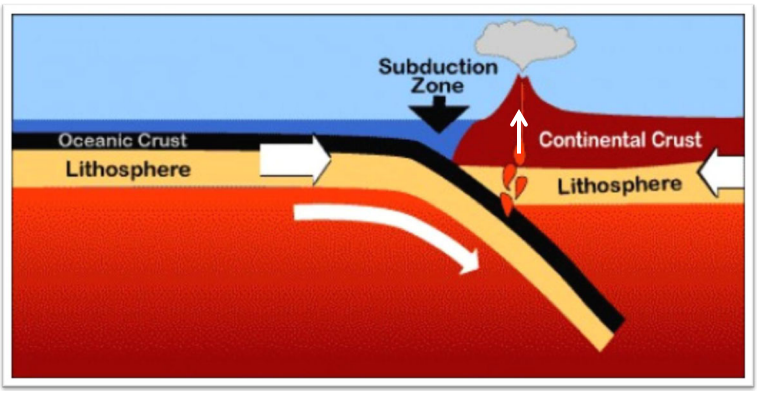

Fig. 1 Example of a picture about volcanism showing how the subduction of tectonic plates results in the creation of a volcanoe

This paragraph is primarily about the picture shown in Fig. 1 and how tectonic plates, subduction, melting and so forth are visualized rather than about the content directly, although the picture deals of course with the content.

The question arises whether the two kinds of paragraphs differ in terms of modality effects. As picture-related paragraphs prompt the reader to look at the corresponding picture, one can assume that there is a need to look at the picture which results in a higher split attention and, accordingly, a stronger modality effect. One can further assume that learners are more likely to follow the prompt if the picture is new to them, whereas they are less likely to do so if they have seen the picture before. These considerations lead to the following hypotheses:

- Picture-related paragraphs are associated with stronger modality effects than content-related paragraphs.
- Modality effects of picture-related paragraphs are stronger if the picture is new to the learner than if the learner has seen the picture before.

In order to test these hypotheses, an experiment on learning from text and pictures was conducted with content-related and picture-related text paragraphs presented in different combinations of visual and auditory modality.

Method

Participants

120 students at grades 11 and 12 from a German gymnasium (an advanced level grammar school) and first year university students participated in the experiment; 47 % were male and 53 % were female. 90 % of the participants were between 17 and 21 years old with an average age of 18.8 years (SD=3.4).

Learning material

A geography textbook about plate tectonics was used as the learning material. The text consisted of a short 81 words introduction into the theory of plate tectonics followed by eight thematic segments dealing with (1) the structure of the Earth's crust, (2) convection currents, (3) types of tectonic movements, (4) convergence of plates, (5) types of magma composition, (6) types of volcanic eruptions, (7) divergence of plates, and (8) rift volcanoes. Total text length (including the introduction) was 1,797 words. Each segment consisted of a content-related paragraph, a picture-related paragraph, and an associated picture. The content-related paragraphs included a total of 1,115 words; the picture-related paragraphs included a total of 601 words. A detailed account of number of words of the paragraphs can be found in Table 3 in the appendix. For each text segment, the readability of the content-related paragraph and the picture-related paragraph were computed according to Amstad (1978). The corresponding scale ranges from 0 (lowest readability) to 100 (highest readability). Readability of the learning material ranged from 29 to 61 with a median of 51. The readabilities of each paragraph are presented in Table 3 in the appendix. The size of the pictures varied between 189 and 420 sq cm with a median of 260 sq cm. The distance between the picture-related text and the picture varied between 1.0 cm and 4.0 cm with a median of 1.4 cm.

Four experimental conditions were used for presenting the text material corresponding to a 2x2 design with the factors *modality of content-related paragraphs (auditory/visual)* and *modality of picture-related paragraphs (auditory/visual). Thus*, for each of the four conditions,

all eight content-related paragraphs of each text were presented either via the visual or the auditory modality. Similarly, all eight picture-related paragraphs were presented either via the visual or the auditory modality. Thus, the 2x2 design for presenting the text material resulted in four experimental conditions: (a) visual content-related text and visual picture-related text, (b) visual content-related text and auditory picture-related text, (c) auditory content-related text and visual picture-related text, (d) auditory content-related text and auditory picture-related text. Written text paragraphs were always presented on the left half of a computer screen, whereas pictures were always presented on the right half of the computer screen. Spoken text paragraphs were presented by a male speaker at a speaking rate of about 100 words per minute through headphones.

According to the original text book, the content-related paragraph preceded the picture-related paragraph within each text segment. This not only corresponds to a widespread instructional practice but also allowed manipulating the novelty of the picture while reading the picture-related paragraph by introducing a third factor in addition to *modality of content-related paragraphs* and *modality of picture-related paragraphs* into the design, namely *novelty of the picture* while processing picture-related paragraphs. That is, on top of the four variants of presenting the text material, two further experimental conditions were introduced. Under the first condition, the picture associated with a text segment was already shown from the outset of the text segment. Thus, for each of the eight text segments, the content-related paragraph was presented with the corresponding picture and the following picture-related paragraph was also presented with the same picture. Under the second condition, the picture associated with a text segment was only shown together with the corresponding picture-related paragraph. In this case, for each of the eight text segments, the content-related paragraph was presented without picture and the following picture-related paragraph was presented with the picture. Picture-novelty while reading the picture-related paragraph was assumed to be lower under the first condition, because the learner had already seen the picture before. Picture-novelty was assumed to be higher under the second condition, because the learner had not seen the picture beforehand.

Figure 2 displays the design in a visual format. Both for the low and for the high picture novelty condition, the content-related paragraphs and the picture-related paragraphs were presented in all possible combinations of auditory and visual modalites. Thus, on the whole, the study used a 2x2x2 design with the factors *modality of content-related paragraphs, modality of picture-related paragraphs* and *novelty of the picture while processing the picture-related paragraphs*.

Instruments

Participants' prior knowledge was assessed based on 20 short statements that had to be judged either as true or false. An example of a prior knowledge statement is 'The Earth's continents are static objects. That is, they do not move.' Learning results were assessed with a post-test consisting of 26 items that did not overlap with the prior-knowledge items. The items of the post-test aimed at comprehension of what was presented in the learning material. They consisted of statements that required participants to draw inferences about the learning content. As an example, students had to judge the statement 'The more water and the more silicic acid the magma contains, the higher is the likelihood of an effusive volcanic eruption' as correct or incorrect and give reasons for their answer.

As students are generally not asked separately about their knowledge acquired from text and knowledge acquired from pictures, there were no separate measures for text-based or picture-based learning. In other words, content-related paragraphs and

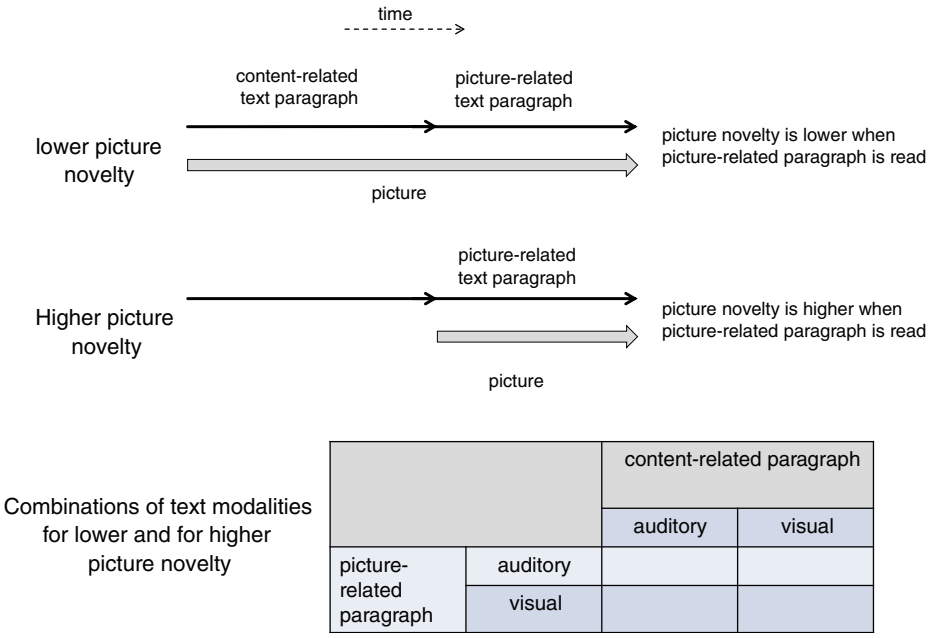

Fig. 2 Visualization of the experimental design. Within each of the eight text segments, the content-related paragraph preceeded the picture-related paragraph. The picture was either shown throughout the text segment (which implied lower novelty when the picture-related paragraph was processed) or it was shown only combined with the picture-related paragraph (which implied higher novelty). All 2x2 combinations of paragraph modalities were implemented for both picture novelty conditions

picture-related paragraphs did not constitute different pieces of knowledge to be tested separately.

Procedure

The 2x2x2-design of the experiment included the following factors: *modality of the content-related paragraphs* (*visual* versus *auditory*), *modality of the picture-related paragraphs* (*visual* versus *auditory*) and *picture novelty* (*low* versus *high*) while reading the picture-related paragraphs. Other factors such as semantic complexity, syntactic complexity, readability scores, text length and picture complexity were controlled in the study by using exactly the same texts and the same pictures in all experimental conditions.

Participants were first tested for their prior knowledge. They were then randomly assigned to one out of the 2x2x2 experimental conditions and then presented with the corresponding variant of the learning material. Instead of processing the learning material under severe time restrictions, students had the possibility to compensate for disadvantages of split attention by sufficient study time. Learning was self-paced within limits. For auditory paragraphs, learning time was defined by the rate of speaking. After listening to a paragraph, participants could proceed to the next paragraph on their own by pressing a "Next" button. If they did not press the button, the presentation of the next paragraph started in a system-based manner after 10 seconds. The pause of 10 seconds was expected to provide some time for finalization of cognitive processing. The maximum time for processing a visual paragraph was exactly the same as for the corresponding auditory paragraph. After reading a paragraph, participants

could proceed to the next paragraph on their own by pressing a "Next" button. If they did not press the button, the presentation of the next paragraph started in a system-based manner after a total presentation time corresponding to the total maximum time of the corresponding auditory paragraph. The maximum time intervals for processing the different paragraphs are presented in Table 3 in the appendix.

Average readers at the age of 17 years and above can be expected to read German texts at a rate of 210 and 250 words per minute (Landerl 2001). As this rate is much higher than a speaking rate of 100 words per minute, participants had enough time for reading the text paragraphs and for observing the pictures without hastiness. Thus, learners could compensate for disadvantages of split attention by using additional study time for the pictures. After learning, students were administered the post-test for assessment of their learning results and received 10 Euro for their participation.

Results

Prior knowledge

Participants' average prior knowledge score in terms of proportion of correctly answered items in the pre-test was 0.74 with a standard deviation of 0.21. The Kolmogorov-Smirnov test revealed a *z*-value of 2.24 indicating a deviation from the normal distribution (p=.001) due to a left skewness of −0.72. This deviation has to be considered in the light of a relatively large sample size which can make also small deviations statistically significant. An inspection of the histograms showed that the deviation was not very strong. Because the sample size was rather large, which simultaneously allows for relatively robust F-values, we considered computation of an ANOVA as acceptable. The 2x2x2 ANOVA of prior knowledge with the factors *content paragraph modality*, *picture paragraph modality* and *picture novelty* revealed the following effects: *content paragraph modality* $F(1,110)=1.48$, $p=.29$, $\eta^2=.010$); *picture paragraph modality* $F(1,110)=2.72$, $p=.10$, $\eta^2=.024$; *content x picture paragraph modality* $F(1,110)=0.12$, $p=.73$, $\eta^2=.001$; *content paragraph modality x picture novelty* $F(1,110)=0.39$, $p=.53$, $\eta^2=.004$; *picture paragraph modality x picture novelty* $F(1,110)=0.48$, $p=.49$, $\eta^2=.004$; *content x picture paragraph modality x picture novelty* $F(1,110)=1.07$, $p=.30$, $\eta^2=.010$. None of these effects was significant. Furthermore, the correlation between prior knowledge and the post-test scores was rather low (r=.11). This is not too surprising because the prior knowledge test aimed at relatively general knowledge, whereas the learning content was rather specific. Due to the lack of group differences and due to the low predictive value of prior knowledge for the post-test scores, there was no need to use prior knowledge as a control variable for an ANCOVA of the post-test results.

Learning time

Participants' average learning time was 22.1 minutes (SD=3.4 min) with a minimum of 17 and a maximum of 33 minutes. The means and standard deviations of the participants' learning time are presented in Table 1. The Kolmogorov-Smirnov test revealed a *z*-value of 1.60 indicating a deviation from the normal distribution (p=.012) due to a right skewness of 1.37. This deviation has to be considered in the light of the relatively large sample size which can make also small deviations statistically significant. An inspection of the histograms showed that the deviation was not strong. Because the sample size was rather large, computation of an ANOVA was considered as acceptable. The 2x2x2 ANOVA of learning time with

Table 1 Means and standard deviations of learning time in minutes

	Modality of content-related paragraphs								
	Visual			Auditory			Total		
	N	M	SD	N	M	SD	N	M	SD
Low picture novelty									
Modality of picture-related paragraphs									
Visual	16	23.5	4.4	15	22.6	3.0	31	23.1	3.7
Auditory	15	22.2	3.2	14	20.6	1.9	29	21.4	2.7
Group total	31	22.9	3.9	29	21.7	2.6	60	22.3	3.4
High picture novelty									
Modality of picture-related paragraphs									
Visual	14	22.6	5.5	15	22.0	2.2	29	22.3	4.1
Auditory	15	21.8	2.4	14	20.7	2.6	29	21.3	2.5
Group total	29	22.2	4.2	29	21.4	2.5	58	21.8	3.5
High and low picture novelty									
Modality of picture-related paragraphs									
Visual	30	23.1	4.8	30	22.3	2.6	60	22.7	3.9
Auditory	30	22.0	2.9	28	20.7	2.2	58	21.4	2.6
Total	60	22.6	4.0	58	21.5	2.5	118	22.1	3.4

the factors *content paragraph modality*, *picture paragraph modality* and *picture novelty* revealed a significant effect picture paragraph modality, F (1,110)=4.30; p=.041; η^2=.040). Learners who had read written picture-related paragraphs needed on average a learning time of 22.7 minutes (SD=3.9) which was 1.3 minutes longer than the average time of 21.4 (SD=2.6) needed by learners who had listened to auditory picture paragraphs. All other effects were not significant. For *content paragraph modality*, the results were F(1,110)=2.68, p=.105, η^2=.026; for *picture novelty* F (1,110)=0.50; p=.48; η^2=.005; *content paragraph x picture paragraph modality* F (1,110)=0.22; p=.64; η^2=.002); *content paragraph modality x picture novelty* F (1,110)=0.07; p=.80; η^2=.001); *picture paragraph modality x picture novelty* F (1,110)=0.19; p=.66; η^2=.002); *content paragraph x picture paragraph modality x picture novelty* F (1,110)=0.00; p=.97; η^2=.000). The Pearson correlation between *learning time* and *post-test scores* was not significant, r=.012.

Learning performance

Participants' post-test scores in terms of the proportion of correctly solved items after learning were distributed normally with a grand mean of 0.67 and a standard deviation of 0.12 with a Kolmogorov-Smirnov z-value of 1.00 below significance (p=.27). The means and standard deviations of the participants' post-test performance are presented in Table 2. As mentioned above, only picture paragraph modality had a significant effect on learning time. As learning time and post-test scores did not significantly correlate, there was no reason to include learning time as a covariate into further analyses. The 2x2x2 ANOVA of the post-test scores with the factors *content paragraph modality*, *picture paragraph modality* and *picture novelty* revealed on the one hand a significant effect of picture paragraph modality, F(1,110)=4.81; p=.015; η^2=.042). Learners who had listened to auditory picture paragraphs showed a significantly

Table 2 Means and standard deviations of post-test test performance (proportion of correctly solved items)

	Modality of content-related paragraphs								
	Visual			Auditory			Total		
	N	M	SD	N	M	SD	N	M	SD
Low picture novelty									
Modality of picture-related paragraphs									
Visual	16	.67	.12	15	.64	.12	31	.65	.12
Auditory	15	.67	.08	14	.65	.10	29	.66	.09
Group total	31	.67	.11	29	.64	.11	.60	.66	.11
High picture novelty									
Modality of picture-related paragraphs									
Visual	14	.66	.14	15	.63	.13	29	.64	.13
Auditory	15	.74	.08	14	.71	.13	29	.73	.11
Group total	29	.70	.12	29	.67	.14	58	.68	.13
High and low picture novelty									
Modality of picture-related paragraphs									
Visual	30	.66	.13	30	.63	.12	60	.65	.13
Auditory	30	.71	.09	28	.68	.12	58	.69	.11
Total	60	.68	.11	58	.66	.12	118	.67	.12

higher mean performance of 0.69 (SD=0.11) than learners who had read written picture paragraphs with a mean performance of 0.65 (SD=0.13). On the other hand, there was no significant effect of content paragraph modality, $F(1,110)=1.46$; $p=.12$; $\eta^2=.013$. As we had hypothesized, picture-related paragraphs led to stronger modality effects than content-related paragraphs. The comparison of the two variances yielded a highly significant difference in favor of the hypothesis, $F(110, 110)=3.23$; $p<.01$.

Furthermore, the 2x2x2 ANOVA revealed a significant interaction *picture paragraph modality x picture novelty,* $F(1,110)=3.41$; $p=.034$; $\eta^2=.030$. In other words, the effect of picture paragraph modality on post-test performance was moderated by picture novelty. The significant interaction *picture-related paragraphs modality x picture novelty* corresponds to our assumption that modality effects of picture-related paragraphs are stronger if the picture is new to the learner. In fact, when picture novelty was high, participants showed a mean performance of 0.73 (SD=0.11) after listening to the picture paragraphs which was significantly higher than their mean performance of only 0.64 (SD=0.13) after reading these paragraphs, $t(56)=2.74$; $p=.004$. On the contrary, when picture novelty was low, participants' mean performance of 0.66 (SD=0.09) after listening and their mean performance of 0.65 (SD= 0.12) after reading the picture paragraphs were nearly the same, $t(56)=0.26$; $p=.40$. All other effects, including the main effect of *picture novelty* ($F(1,110)=1.79$; $p=.18$; $\eta^2=.016$), the interaction *content paragraph modality x picture novelty* ($F(1,110)=0.04$; $p=.85$; $\eta^2=.000$), the interaction *content paragraph modality x picture paragraph modality* ($F(1,110)=0.02$; $p=.88$; $\eta^2=.000$) and the interaction *content paragraph modality x picture paragraph modality x picture novelty* ($F(1,110)=0.01$; $p=.94$; $\eta^2=.030$) were not significant.

To summarize, there was a significant modality effect for the picture-related paragraphs, whereas no modality effect was found for the content-related paragraphs. Accordingly, as we had expected, picture-related paragraphs led to stronger modality effects than content-related

paragraphs. However, the modality effect of picture-related paragraphs appeared only when picture novelty was high, whereas the effect disappeared when the picture had already been seen before. It should be emphasized that picture novelty had only a minor and non-significant overall effect. It obviously did not make a difference whether the picture was presented right from the beginning or only with a picture-related paragraph somewhat later.

For further analysis, a separate ANOVAs for the low and for high picture novelty group was performed. For low picture novelty, the corresponding 2x2 ANOVA revealed neither a significant main effect of *picture paragraph modality* ($F(1,56)=0.07$; $p=.40$; $\eta^2=.001$) nor a significant main effect of *content paragraph modality* ($F(1,56)=0.60$; $p=.22$; $\eta^2=.011$) nor a significant interaction *content x picture paragraph modality* ($F(1,56)=0.03$; $p=.87$; $\eta^2=.001$). For high picture novelty, on the contrary, the corresponding ANOVA revealed a highly significant main effect of *picture paragraph modality* ($F(1,54)=7.19$; $p=.005$; $\eta^2=.12$). The two remaining effects - the main effect of *content paragraph modality* ($F(1,54)=0.86$; $p=.18$; $\eta^2=.016$) and the interaction *content x picture paragraph modality* ($F(1,54)=0.00$; $p=.96$; $\eta^2=.000$) - were not significant.

It was hypothesized that picture-related paragraphs lead to stronger modality effects than content-related paragraphs. Accordingly, we compared these effects under the condition of low and under the condition of high picture novelty. For low picture novelty, the comparison yielded a non-significant difference of $F(54,54)=0.114$. For high picture novelty, on the contrary, the comparison led to a $F(54,54)$ value of 8.385 ($p<.001$). Thus, the effects were significantly different in favor of the hypothesis.

To summarize, there was neither a modality effect of picture-related paragraphs nor a modality effect of content-related paragraphs when pictures were presented right from the beginning. On the contrary, when picture novelty was high (i.e. when pictures had not been seen before), there was a strong modality effect of picture-related paragraphs, but again no modality effect of content-related paragraphs. The lack of a content-related paragraph modality effect when pictures were shown only later has implications for the baseline difference between reading and listening: In the absence of pictures while processing the content-related paragraphs, it did not make an essential difference whether these paragraphs were read or listened to, because the level of performance was essentially the same. In other words, there was no inherent superiority of auditory or visual text presentation under the condition of no picture being presented. Accordingly, the significant effect of picture paragraph modality cannot be attributed to a general superiority of the auditory mode compared to the visual mode because for content-related paragraphs, modality was unrelated to learning performance.

Discussion

The present study aimed at investigating under which conditions a modality effect will occur when learning time is not heavily constrained. Learning time for reading the illustrated text was set equal to the learning time for listening to the illustrated text. Because skilled readers can generally read a text faster than the text can be listened to (or read aloud, respectively), learners might have had sufficient time to investigate the accompanying picture. Students regardless of age frequently process pictures only superficially, because they assume that a short view is enough to grasp the meaning (Hannus and Hyönä 1999; Mokros and Tinker 1987; McDonald and Thornley 2002; Weidenmann 1989). Therefore, instructional design has to adopt measures to enhance sufficiently deep picture comprehension (Bernard 1990; Kulhavy et al. 1985). One possible measure is to provide conceptual guidance for picture processing by verbal descriptions that make the semantic content of pictures explicit.

If such verbal descriptions that provide conceptual guidance for picture processing are included in the learning material, a distinction can be made between different types of text paragraphs. In text linguistics, a distinction is made within sentences or paragraphs between two informational components: the topic and the comment. The topic-component provides (general) information of what the sentence or paragraph is about, whereas the comment-component provides additional (specific) information about this topic (Chafe 1994; Clements 1979; Engelkamp 1982; Givón 1983; Grimes 1975; Halliday 1970; Haviland and Clark 1974). In the study presented above, the topic of content-related paragraphs was volcanism, whereas the topic of picture-related paragraphs was the visualization of volcanism. According to the results presented above, this linguistic distinction makes a difference with regard to the modality effect and its role in instructional design.

There was no modality effect with the content-related paragraphs: One can assume that readers had the opportunity to compensate for the required split attention by investing some additional learning time before proceeding to the next paragraph. This corresponds to the finding that the modality effect disappears or even reverses under self-paced studying conditions (Tabbers et al. 2004; Ginns 2005). Another reason for the lack of a modality effect could be the inherent advantage of reading as compared to listening: the learner's control on his/her pace of processing. Written text is usually stable, which provides more control of cognitive processing, because it allows re-reading difficult passages, whereas spoken text is transitory.

Regarding the picture-related paragraphs, we *did* find a modality effect. However, this effect occurred only under specific conditions. A modality effect for the picture-related paragraphs was found when the pictures were in a strict sense *new* to the learner, that is, when the learner had *not* seen the pictures before. As these paragraphs explicitly refer to the corresponding picture, the learner needs to look at the picture in order to process the paragraph appropriately. Accordingly, there is a strong need for split attention between text and picture under these conditions which results in a significant modality effect. If, however, the picture has already been presented before with the preceding content-related paragraph, the picture is no longer new to the learner and, accordingly, no modality effect is to be expected. In this case, it is possible that the learner already has the required pictorial information in working memory due to previous processing. Thus, there is no further need any more to look at the picture while he/she processes the corresponding picture-related paragraph.

The present investigation used specific material with specific learners. To what extent the present findings can be generalized to other content with other complexity, other text-lengths with other readabilities and other groups of learners of course needs to be subject of further research. Future research should also investigate by eye-tracking methodology whether there is really more split attention between picture-related visual paragraphs and pictures than split attention between content-related visual paragraphs and pictures. Further research should also investigate modality effects when pictures with picture-related paragraphs are presented right at the beginning of the learning material. This is often needed in physics, for example. In this case, the picture-related paragraph precedes the content-related one. If the findings presented above can be generalized, there should be stronger modality effects for picture-related paragraphs than for content-related paragraphs.

As mentioned above, Leahy and Sweller (2011) found a modality effect only when text length was relatively short assuming that this differential effect was due to the transient nature of auditory text. It might be of special interest to further investigate how the modality effect is moderated by picture complexity and text length when a distinction is made between content-related and picture-related paragraphs. Picture complexity and length of picture-related para-

graphs are generally correlated: A complex picture is more difficult to be held in working memory and, thus, needs more glances at the picture for updating its representation in working-memory than a simple picture. A complex picture also needs more verbal guidance and explanation and, thus, results in a longer and more complex picture-related paragraph than a simple picture. If the picture-related paragraphs are presented in the visual format, complex pictures combined with longer picture-related paragraphs might be associated with a stronger need for split attention than simple pictures with shorter paragraphs. Because the auditory modality avoids split attention, the modality effect with picture-related paragraphs should increase rather than decrease with the length of these paragraphs.

As mentioned above previously, the modality effect is not a unitary phenomenon. It seems to be caused by various mechanisms and affected by various factors. The present study suggests that the topic of text paragraphs (i.e. whether they are content-related or picture-related) is another factor which influences whether a modality effect will or will not appear. If the findings of this study can be generalized, the following instructional recommendations can be made. If a high modality effect is to be expected, the corresponding text should be presented in auditory rather than visual modality. A modality effect is to be expected to the extent of how much split attention is required between visual text and pictures. The need for split attention is especially high for picture-related paragraphs which suggests to present picture-related paragraphs generally in the auditory modality. This is especially true if pictures have not been seen before by the learner. Whether content-related paragraphs should also be presented in the auditory or the visual modality depends on how much split attention is required. If the picture is relatively simple or if learners have relatively high prior knowledge or if the density of semantic connections between text and pictures is relatively low, there is no reason to assume a strong need for split attention. In this case, visual text could be more advantageous because it allows for better control of cognitive processing by the learner. Further research is needed to clarify these issues.

Appendix

Table 3 Number of words, maximum processing time, and readability scores of the content-related and the picture-related paragraphs in the experiment

Thematic Text-Segment	Content-Related Paragraphs			Picture-Related Paragraphs		
	No. of Words	Time (sec)	Reada bility	No. of Words	Time (sec)	Readability
Introduction	81	59	29	-		
Structure of the Earth's Crust	169	111	52	82	59	54
Convection Currents	52	40	30	67	50	48
Types of Tectonic Movements	90	62	36	76	56	52
Convergence of Plates	250	158	41	100	70	61
Types of Magma Composition	129	87	39	66	50	52
Types of Volcanic Eruptions	130	88	52	82	59	49
Divergence of Plates	158	105	49	53	42	61
Rift Volcanoes	137	92	61	75	55	51
	1,115			601		

References

Amstad, T. (1978). *Wie verständlich sind unsere Zeitungen.* Dissertation. Universität Zürich.

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedial learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 135–146). Cambridge: Cambridge University Press.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge: Cambridge Univ. Press.

Bernard, R. M. (1990). Using extended captions to improve learning from instructional text. *British Journal of Educational Technology, 21*(3), 215–225.

Chafe, W. L. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Clements, P. (1979). The effects of staging on recall from prose. In R. O. Freedle (Ed.), *New directions in discourse processing. Vol. II in the series Advances in discourse processes* (pp. 287–330). Norwood: Ablex.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

Danks, J. H. (1980). Comprehension in listening and reading: Same or different? In J. Danks & K. Pezdek (Eds.), *Reading and understanding* (pp. 25–40). Newark, Delware: International Reading Association.

Danks, J. H., & End, L. J. (1987). Processing strategies for reading and listening. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 271–294). San Diego: Academic Press.

Diakidoy, I. A., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2012). The relationship between listening and reading comprehension of different types of text at increasing grade levels. *Reading Psychology.* doi:10.1080/02702710590910584.

Ellis, A. W., & Young, A. W. (1996). *Human cognitive neuropsychology. A textbook with readings*. Hove: Psychology Press.

Engelkamp, J. (1982). Given and new information: Theoretical positions and empirical evidence. Saarbrücken: Universität des Saarlandes, Fachrichtung Psychologie im Fachbereich 6, Report No.79.

Ferreira, F., & Anes, M. (1994). Why study spoken language? In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 34–36). San Diego: Academic Press.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313–331. doi:10.1016/j.learninstruc.2005.07.001.

Givón, T. (Ed.). (1983). *Topic continuity in discourse: A quantitative cross-language study.* Amsterdam: Benjamins.

Green, R. (1981). Remembering ideas form text: the effect of modality of presentation. *British Journal of Educational Psychology, 51*, 83–89.

Grimes, J. E. (1975). *The thread of discourse*. The Hague: Mouton.

Gyselinck, V., Jamet, E., & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology, 22*, 353–374.

Halliday, M. A. K. (1970). Language structure and language function. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 140–164). Harmondsworth: Penguin.

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.

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*, 512–521.

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

Hron, A., Kurbjuhn, I., Mandl, H., & Schnotz, W. (1985). Structural inferences in reading and listening. In G. Rickheit & H. Strohner (Eds.), *Inferences in text processing* (pp. 221–245). Amsterdam: North-Holland.

Jeung, H.-J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17*, 329–343. doi:10.1080/0144341970170307.

Kalyuga, S. (2000). When using sound with a text or picture is not beneficial for learning. *Australian Journal of Educational Technology, 16*(2), 161–172.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40*(1), 1–17.

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

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*, 205–226.
Kulhavy, R. W., Lee, B. J., & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology, 10*, 28–37.
Kürschner, C., & Schnotz, W. (2008). Das Verhältnis gesprochener und geschriebener Sprache bei der Konstruktion mentaler Repräsentationen. *Psychologische Rundschau, 59*(3), 139–149.
Landerl, K. (2001). Lesegeschwindigkeitstest (national und international). In G. Haider & B. Lang (Eds.), *PISA PLUS 2000* (pp. 119–130). Studien Verlag: Innsbruck.
Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*(6), 943–951.
Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology, 17*(4), 401–418. doi:10.1002/Acp.877.
Levie, H. W., & Lentz, R. (1982). Effects of text illustrations: a review of research. *Educational Communication and Technology Journal, 30*, 195–232.
Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In H. A. Houghton & D. M. Willows (Eds.), *The psychology of illustration* (pp. 51–86). New York: Springer.
Lindow, S., Fuchs, H. M., Fürstenberg, A., Kleber, J., Schweppe, J., & Rummer, R. (2011). On the robustness of the modality effect: Attempting to replicate a basic finding. *Zeitschrift für Pädagogische Psychologie, 25*(4), 231–243.
Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 147–158). Cambridge: Cambridge University Press.
Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225–244.
Lowe, R. K. (2004). Interrogation of a dynamic visualisation during learning. *Learning and Instruction, 14*, 257–274.
Massaro, D. W. (1979). Reading and listening. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (pp. 331–370). New York: Plenum Press.
Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational Psychologist, 32*, 1–19.
Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge: Cambridge University Press.
Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.
Mayer, R. E., & Moreno, R. (1998). Split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*, 312–320.
Mayrath, M. C., Nihalani, P. K., & Robinson, D. H. (2011). Varying tutorial modality and interface restrictions to maximize transfer in a complex simulation environment. *Journal of Educational Psychology, 103*, 257–268.
McDonald, T., & Thornley, C. (2002). Unlocking meaning in text: some thoughts on the literacy challenges our students face. *English in Aotearoa, 48*, 54–60.
Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity to process information. *Psychological Review, 63*, 81–97.
Mokros, J. R., & Tinker, R. F. (1987). The Impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching, 24*, 369–383.
Moreno, R., & Durán, R. (2004). Do multiple representations need explanations? The role of verbal guidance and individual differences in multimedia mathematics learning. *Journal of Educational Psychology, 96*(3), 492–503.
Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology, 91*, 358–368.
Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319–334. doi:10.1037/0022-0663.87.2.319.
Peterson, L., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193–198.
Ploetzner, R., Lowe, R., & Schlag, S. (2013). A systematic characterization of cognitive techniques for learning from textual and pictorial representations. *Journal of Education and Learning, 2*(2), 78–95.
Rieben, L., & Perfetti, C. (1991). *Learning to read: Basic research and its implications*. Hillsdale: Erlbaum.
Rubin, D. L., Hafer, T., & Arata, K. (2000). Reading and listening to oral-based versus literate-based discourse. *Communication Education, 49*(2), 121–133.
Rummer, R., Schweppe, J., Fürstenberg, A., Seufert, T., & Brünken, R. (2010). Working memory interference during processing texts and pictures: implications for the explanation of the modality effect. *Applied Cognitive Psychology, 24*, 164–176.
Schmidt-Weigand, F. (2011). Does animation amplify the modality effect – or is there any modality effect at all? *Zeitschrift für Pädagogische Psychologie.*

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010a). Explaining the modality and contiguity effects: New insights from investigating Students' viewing behaviour. *Applied Cognitive Psychology, 24*, 226–237. doi:10.1002/Acp.1554.

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010b). 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.

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

Schnotz, W. (2011). Colorful bouquets in multimedia research: a closer look at the modality effect. *Zeitschrift für Pädagogische Psychologie, 25*, 269–276.

Schüler, A., Scheiter, K., & Schmidt-Weigand, F. (2011). Boundary conditions and constraints of the modality effect. *Zeitschrift für Pädagogische Psychologie, 25*, 211–220.

Sticht, T. G., & James, J. H. (1984). Listening and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 293–317). New York: Longmann.

Stiller, K. D., Freitag, A., Zinnbauer, P., & Freitag, C. (2009). How pacing of multimedia instructions can influence modality effects: a case of superiority of visual texts. *Australasian Journal of Educational Technology, 25*, 184–203.

Tabbers, H. K., & van der Spoel, W. (2011). Where did the modality principle in multimedia learning go? A double replication failure that questions both theory and practical use. *Zeitschrift für Pädagogische Psychologie, 25*(4), 221–230.

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

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

Weidenmann, B. (1989). When good pictures fail: An information-processing approach to the effects of illustrations. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 157–170). Amsterdam: Elsevier.

Yeung, A. S., Jin, P., & Sweller, J. (1997). Cognitive load and learner expertise: split-attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology, 23*, 1–21.

Wolfgang Schnotz. Faculty of Psychology, General and Educational Psychology, University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany. schnotz@uni-landau.de. Phone: +49-634128034234; Fax: +49-634128034240; Email: schnotz@uni-landau.de

Current themes of research:

Text-Picture Integration, Knowledge Acquisition from Multiple Representations, Knowledge Acquisition with Conflicting Information, Multimedia Learning, Learning with Hypermedia, Learning with Animation

Most relevant publications in the field of Psychology of Education:

Schnotz, W. (2011). Colorful Bouquets in Multimedia Research: A Closer Look at the Modality Effect. Zeitschrift für Pädagogische Psychologie. 25, 269–276.

Schnotz, W., Baadte, C., Müller, A., & Rasch, R. (2010). Creative thinking and problem solving with depictive and descriptive representations. . In L. Verschaffel, E. De Corte, T. de Jong & J. Elen (Eds.) Use of representations in reasoning and problem solving (pp. 11–35). London: Routledge.

Schnotz, W., & Heiß, A. (2009). Semantic Scaffolds in Hypermedia Learning Environments. Computers in Human Behavior, 25(2), 371–380.

Schnotz, W. & Kürschner, C. (2007). A reconsideration of cognitive load theory. Educational Psychology Review, 19(4), 469–508.

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

Christoph Mengelkamp. University of Würzburg, Philosophical Faculty II, Oswald-Külpe-Weg 82, 97074 Würzburg, Germany. christoph.mengelkamp@uni-wuerzburg.de

Current themes of research:

metacognition in self-regulated learning, interactive animations, learning from pictures and texts, educational diagnostics and methodology

Most relevant publications in the field of Psychology of Education:

Mengelkamp, C., & Bannert, M. (2012). Confidence judgments in learning. In N. M. Seel (Ed.), Encyclopledia of the sciences of learning (pp. 756–759). New York: Springer. doi:10.1007/978-1-4419-1428-6 doi:10.1007/978-1-4419-1428-6#_blank

Mengelkamp, C., & Bannert, M. (2010). *Accuracy of confidence judgments: Stability and generality in the learning process and predictive validity for learning outcome.* Memory and Cognition, 38, 441–451. doi:10.3758/MC.38.4.441 doi:10.3758/MC.38.4.441#_blank

Bannert, M., Hildebrand, M., & Mengelkamp, C. (2009). *Effects of a metacognitive support device in learning environments.* Computers in Human Behaviour, 25, 829–835. doi:10.1016/j.chb.2008.07.002 doi:10.1016/j.chb.2008.07.002#_blank

Mengelkamp, C., & Bannert, M. (2009). *Judgements about knowledge: searching for factors that influence their validity.* Electronic Journal of Research in Educational Psychology, 7(1), 163–190.

Bannert, M., & Mengelkamp, C. (2008). *Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning?* Metacognition and Learning, 3, 39–58. doi:10.1007/s11409-007-9009-6 doi:10.1007/s11409-007-9009-6#_blank .

Christiane Baadte. University of Koblenz-Landau, Faculty of Psychology, Fortstraße 7, 76829 Landau, Germany. baadte@uni-landau.de

Current themes of research:

Diagnostic and didactic competencies of teachers, feedback with self-regulated learning with text-picture integration, working memory and multimedia learning

Most relevant publications in the field of Psychology of Education:

Baadte, C. & Dutke, S. (in press). Learning about persons: The effects of text structure and executive capacity on conceptual change. *European Journal of Psychology of Education*. doi:10.1007/s10212-012-0153-2

Baadte, C. & Dutke, S. (in press). Learning About Persons: The Effects of Text Structure and Executive Capacity on Conceptual Change. *European Journal of Psychology of Education*.

Baadte, C. & Schnotz, W. (in press). Feedback Effects on performance, motivation and mood: Are they moderated by the learner's self-concept? *Scandinavian Journal of Educational Research*.

Baadte, C. & Schnotz, W. (2012). Grundlagen des Verstehens von Texten mit Bildern. *Weiterbildung, 6*, S.35-38.

Dutke, S., Baadte, C., Hähnel, A., von Hecker, U., & Rinck, M. (2010). Using diagnostic text information to constrain situation models. *Discourse Processes, 47* (6), 510–544.

Georg Hauck. Rhine-Waal University of Applied Sciences, Communication and Environment, Südstraße 8, 47475 Kamp-Lintfort, Germany. georg.hauck@hochschule-rhein-waal.de

Current themes of research:

Instructional media, communication and organizational counseling

Most relevant publications in the field of Psychology of Education:

Kürschner, C., Seufert, T., Hauck, G., Schnotz, W. & Eid, M. (2006). Konstruktion visuell-räumlicher Repräsentationen beim Hör- und Leseverstehen. *Zeitschrift für Psychologie, 214*(3), S. 117–132.

Kürschner, C., Schnotz, W., Eid, M. & Hauck, G. (2005). Individuelle Modalitätspräferenzen beim Textverstehen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 37*(1), 2–16.

Hauck, G. (2008). *Der Aufbau von Wissensstrukturen mit Bildern*. Saarbrücken: VDM.

Hauck, G. (2005). *Effekte von Modalitätspräferenzen beim Wissenserwerb mit multimedialen Lernsystemen*. Berlin: Logos.