Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content

Detlev Leutner *, Claudia Leopold, Elke Sumfleth

Duisburg-Essen University, Instructional Psychology Dept., Faculty of Education Sciences, P.O. Box D-45117 Essen, Germany

Article history:
Available online 12 January 2009

Keywords:
Cognitive load
Multimedia
Science learning
Text comprehension
Drawing pictures
Mental imagery

ABSTRACT

One hundred and eleven 10th graders read an expository science text on the dipole character of water molecules (ca. 1600 words). Reading instruction was varied according to a 2 × 2 experimental design with factors ‘drawing pictures of text content on paper’ (yes, no) and ‘mentally imagining text content while reading’ (yes, no). The results indicate that drawing pictures, mediated through increased cognitive load, decreased text comprehension and, thus, learning (d = 0.37), whereas mental imagery, although decreasing cognitive load, increased comprehension only when students did not have to draw pictures simultaneously (d = 0.72). No evidence was found that the effects were moderated by domain-specific prior knowledge, verbal ability, or spatial ability. The results are in line with cognitive theories of multimedia learning, self-regulated learning, and mental imagery as well as conceptions of science learning that focus on promoting mental model construction by actively visualizing the content to be learned. Constructing mental images seems to reduce cognitive load and to increase comprehension and learning outcome when the mental visualization processes are not disturbed by externally drawing pictures on paper, whereas drawing pictures seems to increase cognitive load resulting in reduced comprehension and learning outcome.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Recent research on learning with multimedia has revealed that a combination of verbal and pictorial material usually helps learning (the so-called “multimedia effect”; e.g., Mayer, 2001, 2005; Schnitz, 2005). For example, when an expository text is illustrated with suitable pictures which depict the spatial relations of functional elements that are described and discussed in the text (representational visualizations according to Carney & Levin, 2002), students can better comprehend the text – regardless whether it is written on paper or displayed in a hypermedia learning environment.

On the other hand, research on self-regulated learning has revealed that using cognitive and metacognitive strategies also helps learning (Leutner & Leopold, 2006; Pintrich, 2000; Schreiber, 1998; Weinstein & Mayer, 1986). For example, when students have to read an expository text for comprehension and learning, they profit from using cognitive, deep-level learning strategies like text-highlighting or concept mapping (for the latter see, e.g., Hilbert & Renkl, 2009; or Van Gog, Kester, Nieveelstein, Giesbers, & Paas, 2009) as well as from controlling the application of their cognitive strategies by using metacognitive strategies like planning and monitoring (Leopold, den Elzen-Rump, & Leutner, 2007; Leutner, Leopold, & den Elzen-Rump, 2007).

Combining ideas of multimedia learning and self-regulated learning, the question arises how learners can improve their understanding when there are no representational pictures included in an expository text. Do instructions that activate learners to generate and construct pictures by themselves enhance text comprehension and, thus, learning? In that case, the construction of pictures can be regarded as the self-regulated application of a cognitive learning strategy. Thus, constructing pictures for understanding, either by drawing them on paper or by simply visualizing them mentally when reading a text, may be called “self-regulated visualizing”. For example, Leopold et al. (2007) found that drawing pictures as well as mentally imagining pictures (images) of text content was often spontaneously used when students read a science text (see also van Meter, 2001; van Meter & Garner, 2005, concerning drawing strategies; and Cooper, Tindall-Ford, Chandler, & Sweller, 2001; Kosslyn, 1994, for mental imagery strategies).

Learning from expository texts using a self-regulated visualization strategy (either drawing pictures or mentally imagining text content) seems to be a reasonable learning strategy when pictures are helpful for comprehension, but are not provided in the text. However, visualization strategies might impose high demands on the learner in terms of cognitive load (Ainsworth, 1999). According to cognitive load theory (Paas, Renkl, & Sweller, 2003; Sweller,
1994), three types of cognitive load can be conceptually distin-
guished from each other: intrinsic cognitive load, extraneous cog-
nitive load, and germane cognitive load. “Intrinsic” cognitive load is
induced by the complexity of the subject matter, “extraneous”
cognitive load by a deficient instructional design of the learning
material, and “germane” cognitive load by cognitive processes that
lead to deeper understanding of the learning material. All three
types of cognitive load are normally present during learning, and
cognitive overload, inhibiting learning, occurs, when the sum of
the three loads exceeds the limitations of a student’s working
memory. From a cognitive-load perspective, it is not intuitively
clear, which kind of load is induced by using a self-regulated visu-
alization strategy in learning from a text without pictures, that is,
with insufficiently designed instructional material. On the one
hand, the visualization strategy might be necessary – at least for
some learners – in order to compensate bad instructional design
(i.e., text without pictures), thus inducing extraneous cognitive
load. This type of load arises because learners are required to trans-
form verbal information into pictorial information which is other-
wise provided to them in the learning material when pictures are
included. On the other hand, although inducing extraneous load,
the visualization strategies – initiated by the instructional design
of a text without pictures – can be expected to induce cognitive
processes that lead to deeper understanding, thus inducing ger-
mane cognitive load.

Another question concerns the role and the specific effect of
transforming the verbal information, while reading a text, into a
pictorial representation by externally drawing pictures on paper
as opposed to mentally imagining pictures without externally
drawing them on paper. When imagining text content, learners
have to mentally transform verbal information into pictorial infor-
mation which fosters deeper processing. When drawing pictures of
text content on paper, however, learners have to mentally trans-
form verbal information into pictorial information and have to
externalize the pictorial information, which is expected to require
additional cognitive resources. Furthermore, according to a num-
ber of studies (see Zwaan, 2004), generating internal analogue rep-
resentations can be regarded to be a highly automatic process that
facilitates the construction of a mental model and occurs when
learners listen to verbal narrations or when reading a text. Thus,
it can be expected that using a mental imagery strategy will foster
text comprehension without imposing too much cognitive load on
the learner. On the other hand, drawing pictures of text content on
paper is – for most learners – far from being a strategy that builds
upon automatic cognitive processing. In fact, drawing pictures on
paper when reading a text might represent a secondary task that
 imposes additional cognitive load on the learner in general com-
 peting with cognitive resources that are otherwise required for
 accomplishing the primary task of understanding the text and
 building a suitable mental model of the text content. However,
 explicit instruction to externally draw pictures when reading a text
 might help specific learners that are otherwise not able or not used
to construct adequate mental images, for example, due to low do-
main-specific content knowledge or low spatial abilities (see e.g.,
Mayer, 2001, on prior knowledge and spatial ability as moderators
of instructional design). In such cases, concerning visualization
strategies, external pictures, drawn by the learner on paper, might
facilitate metacognitive processes of mental model quality control
that otherwise, when focusing on transient mental images only, are
not triggered at all. Thus, although the instruction to draw pictures
when reading a text might impose additional extraneous cognitive
load at least on some learners, this load might be classified as ger-
mane load for other learners because it can trigger metacognitive
processes that lead to better text comprehension and better learn-
ing (i.e., representing some kind of “expertise reversal” effect,
Kalyuga, 2005, or – more general – a specific type of “aptitude-
treatment interaction”, Cronbach & Snow, 1977, or – even more
 general – “trait–treatment interaction”, Leutner & Rammsayer,
1995).

Thus, it is an open question which role cognitive load plays
when using a visualization strategy while reading a non-fictional,
expository text without pictures. In order to answer this ques-
tion, an experimental study was conducted in which the instruc-
tion on how to proceed when reading an instructional science
text was varied by fostering either internal pictorial representa-
tions (mental imagery) and/or external pictorial representations
(drawing of pictures) of text content. Furthermore, it was inves-
tigated whether the effects of mental imagery and picture draw-
ing on text comprehension are at least to some extent mediated
by cognitive load and/or moderated by domain-specific prior
knowledge, verbal ability, or spatial ability. Due to the explor-
atory nature of the study, however, no specific hypotheses were
formulated.

2. Method

2.1. Participants

One hundred and eleven German students, 10th graders of a
higher track secondary school, participated in the study (56 boys
and 55 girls; age $\text{M} = 16.1$ years, $\text{SD} = 0.44$). Within their classes,
students were randomly assigned to one of four treatment condi-
tions with $27 < N < 29$ students per treatment group. The groups
are balanced according to students’ gender, $\chi^2(3) = 1.24, p = .743$.

2.2. Design and materials

The study followed a $2 \times 2$ experimental design with drawing
instruction (yes, no) and mental imagery instruction (yes, no) as
the two experimental factors. Students had to read a science text
on the dipole character of water (approximately 1600 words;
structured into 13 paragraphs the content of which is depicted in
Fig. 1). According to the specific experimental treatment condition,
students were instructed to read the text and either to mentally
imagine the content of each paragraph of the text while reading
the paragraph, to draw pictures on a sheet of paper representing
the content of each paragraph of the text, or to do both by imagi-
ning the pictures after drawing them. When having to do both, for
example, students were instructed to perform three steps for each
paragraph of the text to be read: “(1) Read the paragraph, (2) draw
a picture that represents the content of the paragraph, and (3) gen-
erate a mental image of your picture!” Students of the three treat-
ment conditions were told that their pictures (or images) should
be simple and clear and that they should represent the most impor-
tant information of the paragraph. The pictures (or images) should
help them to understand the text. Students of the control treat-
ment condition were instructed to read the text for comprehension
only.

The dependent variable was students’ amount of text compre-
prehension after having read the text, measured by a multiple-choice
test on the content of the science text (22 items with 1–4 correct or
false alternatives each; Cronbach’s $\alpha = .84$). Item examples are
“What is the basic principle of a hydrogen bond? (a) The polar
ature of the water molecule, (b) the attraction forces between elec-
trons, (c) the attraction forces between ions, or (d) the polar
covalent bond occurring in the water molecule?” and “During
hydration, a ‘ion-dipol’ attraction is being established. What is
meant with this term? (a) Forces between two or more ions, (b)
forces between water molecules and ions, (c) forces within water
molecules, or (d) forces between the electrons of the ions?” The fo-
cus of the test was on comprehension, not on recall, because the
correct answers to the questions could not be directly taken but had to be inferred from the text.

In order to account for individual differences in domain-specific prior knowledge, a science pre-test was used consisting of 10 multiple-choice items measuring knowledge of basic concepts (e.g., “What is a molecule?”) as well as comprehension of basic principles (e.g., “Why do icebergs float in water?”). As potential moderator variables, that might moderate the effect of the experimental factors on the dependent variable, verbal ability was measured using a standard intelligence test (scale “word fluency”; Heller & Perleth, 2000) as well as spatial ability using a paper-folding test (Ekstrom, French, & Harman, 1976).

As a potential mediator variable that might mediate the effect of the experimental factors on the dependent variable, the amount of cognitive load experienced while reading the science text under the specific reading instruction was measured using a questionnaire with five items, each with a seven-point rating scale (Cronbach’s α = .80). The following items were used: (1) “When reading for comprehension I invested very low...very high mental effort” (r_{it} = .55); (2) “Working on the text represented a strong under-challenge...strong over-challenge” (r_{it} = .70); (3) “After working on the text I felt relaxed and refreshed...very zapped” (r_{it} = .43); (4) “Comprehending the text was very easy...very difficult” (r_{it} = .72); (5) “The topic of the text per se I felt to be very easy...very difficult” (r_{it} = .60). Item (1) represents the often used Paas & van Merriënboer item for measuring cognitive load in terms of invested mental effort (Paas & van Merriënboer, 1993). The other items were included for tapping further aspects of cognitive load that go beyond mental effort, specifically aspects of perceived level of difficulty (see, e.g., for a discussion of this point, Cierniak, Scheiter, & Gerjets, 2009; Schnitz & Kürschner, 2007; or Van Gog & Paas, 2008). However, the high level of internal consistency of the five items and the high (corrected) item-total correlations r_{it} of these items with the scale score indicates that the participants were not able to differentiate these aspects of cognitive load. Thus, the average rating of each participant across the five items was used as a measure of cognitive load in the data analyses below.

2.3. Procedure

Within their classes, students were randomly allocated to one of the four experimental treatment conditions, and each student received a booklet according to the condition he or she was allocated to. At first, students filled in the science pre-test; then they were given 35 min to read the science text with a specific instruction according to their experimental treatment condition. Afterwards they filled in the cognitive-load questionnaire, and – in order to increase the gap between reading the text and testing text comprehension – the verbal-ability test and the spatial-ability test. Last but not least, the multiple-choice post-test on science text comprehension was administered. When working through the post-test, students did not have access to the science text. Overall, the whole procedure took about 90 min.

3. Results

Data were analysed by analyses of variance and covariance within the general-linear-model framework. Moderator effects were tested using the sequential decomposition of variance option of SPSS, starting with the moderator variable as the first predictor in the linear model, then entering the treatment main and interac-
tion effect variables into the model and afterwards, as the last step, entering the interactions of the moderator variable and the treatment effect variables. **Mediator effects** were tested using the procedure proposed by Baron and Kenny (1986). Before looking at treatment effects on the dependent variables, it was ensured that the four experimental groups do not differ regarding domain-specific prior knowledge, verbal ability, and spatial ability, all $F(3,107) < 1$.

Concerning text comprehension, as the main dependent variable, analysis of variance indicates a negative main effect of drawing instruction, $F(1,107) = 3.98$, $MSE = 74.65$, $p = .049$, $d = -.37$, but no main effect of mental imagery instruction, $F(1,107) = 1.04$, $MSE = 74.65$, $p = .310$, $d = .019$. Students that were not instructed to draw pictures representing text paragraph content ($M = 67.0$, $SD = 9.6$) outperformed students that were instructed to draw pictures ($M = 63.7$, $SD = 8.4$). However, the negative main effect of drawing instruction is moderated by the mental imagery instruction as shown by a statistically significant interaction of the two experimental effect variables, $F(1,107) = 9.92$, $MSE = 74.65$, $p = .002$, $r^2 = .09$. The overall pattern of results is depicted in Fig. 2a: when mental imagery instruction is present, the effect of drawing pictures is negative and in particular large ($M = 63.4$, $SD = 9.3$, control condition without drawing instruction; $d = .023$). Thus, compared to the control condition without any specific reading strategy instruction, drawing pictures has some slight advantage that turns into a disadvantage, when both drawing and mental imagery of the drawn picture is requested from the learner. On the other hand, mental imagery instruction has a rather large advantage as far as mental imagery instruction is not combined with drawing instruction. That is, it seems beneficial when learners are explicitly instructed to mentally imagine the content of a text paragraph while reading that paragraph (mental imagery instruction), but it seems not beneficial when they are explicitly instructed to mentally imagine the picture they have drawn while reading the text paragraph (drawing and mental imagery instruction).

Concerning cognitive load (in terms of mental effort and experienced difficulty), significant main effects occurred (Fig. 2b; the load variable being rescaled so that high means indicate low load): drawing instruction increased, $F(1,107) = 14.85$, $MSE = 1.06$, $p < .001$, $d = .72$; whereas mental imagery instruction decreased, $F(1,107) = 5.53$, $MSE = 1.06$, $p = .021$, $d = -.44$, the cognitive load experienced while reading the text, the interaction being not statistically significant, $F(1,107) = 1.34$, $MSE = 1.06$, $p = .245$. With $d = .72$, the negative load effect of drawing instruction is rather large ($M = 4.5$, $SD = 1.1$, without drawing instruction; $M = 3.8$, $SD = 1.0$, with drawing instruction), whereas, with $d = -.44$, the positive load effect of mental imagery instruction is of medium size ($M = 3.9$, $SD = 1.1$, without mental imagery instruction; $M = 4.4$, $SD = 1.1$, with mental imagery instruction). Thus, the drawing instruction loaded whereas the mental imagery instruction off-loaded the demands on working memory while reading the text.

Note that there is no evidence that any of the effects of reading strategy instruction is moderated by domain-specific prior knowledge, verbal fluency, or spatial ability: **mediator analyses** including prior knowledge ($M = 70.4\%$ correct, $SD = 9.6$), verbal fluency ($M = 67.8\%$ correct, $SD = 11.9$) and spatial ability ($M = 65.5\%$ correct, $SD = 18.9$) as covariates in the linear model and testing the statistical interaction of these covariates and treatments, revealed statistically non-significant effects ($0.01 < F(1.103) < 1.71$). The **mediator analysis**, however, shows that the negative main effect of drawing instruction on comprehension is mediated by increased cognitive load. Fig. 3a shows that there are statistically significant effects of drawing instruction on cognitive load and on comprehension as well as a significant effect of cognitive load on comprehension. However, when both variables, drawing instruction and cognitive load, are included in the linear model for predicting comprehen-

---

**Fig. 2.** Mean comprehension and mean cognitive load, as a function of mental imagery and drawing instruction.

**Fig. 3.** Results of mediator analyses.
sion, that is when the effect of drawing instruction is adjusted for cognitive load, the effect of drawing instruction on comprehension is reduced from $\eta^2 = .04$ to $\eta^2 < .01$ which is no longer statistically significant. In other words: There is no direct negative effect of drawing instruction on comprehension; the negative effect of drawing instruction on comprehension is completely mediated by its effect on increased cognitive load ($\eta^2 = .14$) which in turn reduces comprehension ($\eta^2 = .18$). Fig. 3b shows that mental imagery instruction does not have an effect on comprehension so that there can also be no mediation by cognitive load. Fig. 3c shows that the interaction of drawing and mental imagery instruction has an effect on comprehension; this effect, however, is not mediated by cognitive load because there is no significant effect of the interaction on cognitive load. Thus, when adjusting the effects of drawing and mental imagery instruction on comprehension regarding cognitive load, the main effect of drawing instruction vanishes whereas the interaction of drawing and mental imagery instruction does not vanish, $F(1,106) = 8.39$, $MSE = 64.80$, $p = .005$, $\eta^2 = .08$ (Fig. 4).

4. Summary and discussion

The present study investigated whether learning from a non-fictional, expository science text that does not include any pictures can be improved by instructing students to read the text paragraph-wise and either to draw a picture of text content after reading each paragraph and/or to mentally imagine the text content while reading the paragraph. Theoretical background are theories of multimedia learning (e.g., Mayer, 2001, 2005; Schnotz, 2005), that is, learning with verbal and pictorial material in an auditory and/or visual presentation format, on the one hand, and theories and models of self-regulated learning (Leutner & Leopold, 2006; Pintrich, 2000; Weinstein & Mayer, 1986), that is, using cognitive and metacognitive strategies for deep understanding of the learning materials, on the other. Self-regulated learning comes into play because, when reading a text without pictures, students may use visualization learning strategies in order to self-construct missing pictures of text content, either in form of internal, mental images or as external drawings on paper (Denis & Cocude, 1992; Kosslyn, 1994; Leopold et al., 2007; van Meter, 2001; van Meter & Garner, 2005). In other words, it was expected that the advantages of multimedia-based learning with verbal and pictorial material can unfold its impact on learning outcome, even when there is verbal material only and when students are instructed to construct the pictorial material by themselves. However, self-constructing pictures while reading and learning, either mentally or drawn on paper, may pose additional cognitive load on the learner that can, according to cognitive load theory (Paas et al., 2003; Sweller, 1994), turn to be either extraneous or germane and thus either hinder or promote learning.

Concerning cognitive load, the results of the present study indicate that the instruction to draw pictures of text content while reading a science text increases cognitive load while the instruction to mentally imagine text content decreases cognitive load. Furthermore, mediator analyses showed that increased cognitive load due to drawing external pictures impairs reading comprehension and, thus, learning. In other words, what is intended to trigger helpful cognitive processing (drawing pictures for understanding in order to impose germane cognitive load on the learner and to help learning) entailed to impose extraneous cognitive load that hinders learning. On the other hand, decreased cognitive load due to constructing mental images has, in terms of a main effect, no direct impact on reading comprehension and, thus, on learning. The specific combination of drawing and mental imagery instruction, however, affects reading comprehension in terms of a disordinal interaction that turned out not to be mediated by cognitive load. The interaction pattern indicates that constructing mental images promotes reading comprehension and learning only when the mental imagery instruction is focused directly on the text content while reading for comprehension rather than when the mental imagery instruction is focused on a picture drawn from the text in order to externally represent the text content and, thus, when the mental imagery instruction is focused indirectly on the text content. Or in other words: Constructing mental images can only unfold its positive impact on text comprehension and learning when it is performed directly while reading the text and not, as in the present study implemented, when it is focused on an external picture drawn from the text on paper. Of course, in the drawing only condition, before drawing the external picture, a learner can be expected to have a mental image of the text content (like in the imagining only condition). However, the positive effect of this mental image seems to vanish when the image has to be transformed into an external picture on paper.

This result that instructionally induced mental imagery of text content while reading an expository text helps learning whereas instructionally induced picture drawing of text content and mental imagery of the pictures after reading the text impairs learning, is of specific theoretical interest. The result is in line with cognitive theories of mental imagery (Kosslyn, 1994) and, more specifically, theories of generating internal analogue representations as an automatic process of mental model construction while processing verbal information (Schnotz, 2005; Zwaan, 2004). Zwaan (for an overview see Zwaan, 2004), for example, has shown in a number of studies that participants, when listening to words and sentences describing concrete objects, situations, and actions, generate internal analogue representations of those objects, situations, and actions, the perceptual features (Zwaan, Stanfield, & Yaxley, 2002) and psycho-motor affordances (Zwaan & Taylor, 2006) of which correspond to the real objects. This cognitive process of transforming verbal information into internal analogue representations seems to be highly automatic and often leads to a highly consistent and integrated situation or mental model.

The results of the present study indicate that the effectiveness of transforming verbal information into a mental model from which answers to comprehension questions can successfully be derived, can be increased by explicitly instructing students to do so. However, it seems to be of specific interest that the transformation process does not, on the one hand, increase the overall amount of cognitive load and does not, on the other hand, interrupt the flow of processing the verbal information. Mental imagery while reading an expository text, as opposed to picture drawing, seems to have these helpful process characteristics whereas picture drawing does not, leading to an incomplete and/or inconsistent mental

![Fig. 4. Mean comprehension adjusted for cognitive load, as a function of mental imagery and drawing instruction.](image-url)
model and, thus, to less comprehension and reduced learning outcomes. It should be added that this difference in the learning effectiveness of generating mental images and drawing external pictures is in line with the idea of split attention (e.g., Ayres & Sweller, 2005): Drawing a picture, when reading an expository text, requires to switch back and forth between text and picture in order to search for further text information that can be or has to be visualized in the picture. This repeated switching between text and picture, however, requires cognitive resources at the expense of those cognitive processes that can lead to better comprehension and better learning. On the other hand, those highly automated cognitive processes of mentally visualizing text information while reading a text do not seem to take off cognitive resources that can be better used for comprehension and learning.

Of course, in the present study, the advantages of mental imagery over picture drawing as an effective and efficient cognitive strategy for learning from text were found with specific learners, specific learning content, and specific learning materials. Concerning content and materials, the results are based on science learning by self-paced reading an expository text; whether the results hold for other subject matters and other presentation formats is an open question. It can be expected that when other learning content is concrete, describing objects or situation that can be represented by representational pictures, comparable results will be obtained. Whether the presentation format of verbal information (auditory as narration or visually as written text), however, does matter is an open question. Concerning learner characteristics, one has to have in mind, that drawing pictures from text imposed, in the present study, high cognitive load on the learners that hindered comprehension and learning. Again, it is an open question, whether this effect might disappear when the drawing of pictures would not impose so much cognitive load that might, at least in part, be due to students’ low experience in drawing pictures for learning. Studies by Leutner et al. (2007) and Leopold et al. (2007) yielded promising results that students can successfully be trained in visualizing text content (as well as other cognitive learning strategies like text-highlighting or concept mapping) as an effective cognitive learning strategy. However, in order to have a greater impact on the outcome of learning from text when applying the strategy, the trainings should have a specific focus on metacognition, as is shown in the afore-mentioned studies (Leopold et al., 2007; Leutner et al., 2007). That is, when applying a visualization strategy, earners profit from monitoring and controlling each step of the strategy in order to derive at high quality pictures that are indeed accurate and valid with respect to what is written in the text and to be learned.

Acknowledgements

The present publication is based on research projects funded by the German Research Foundation (DFG; LE 645/9-1 and -2). The authors thank Janneh Magdo who helped in collecting the data.

References