Learning and Instruction 38 (2015) 24–33

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

The multimedia effect and its stability over time

Judith Schweppe [a, *], Alexander Eitel [b], Ralf Rummer [a]

[a] University of Erfurt, Germany
[b] Knowledge Media Research Center, Tübingen, Germany

ARTICLE INFO

Article history:
Received 12 August 2014
Received in revised form
6 March 2015
Accepted 9 March 2015
Available online

Keywords:
Multimedia effect
Long-term memory
Mental model
Desirable difficulties

ABSTRACT

A central finding in learning with text and pictures is a benefit from learning with combined representations as compared to text only – the multimedia effect. Two experiments (Experiment 1: N = 110; Experiment 2: N = 147) test for whether this advantage is restricted to short-term learning or whether it remains stable when learning is tested after a delay. Subjects learned about a pulley system at their own pace, either with only a written text or with text and a picture of the described pulley system presented simultaneously. We varied whether comprehension and retention were tested immediately after the learning phase or after a delay (1 week later in Experiment 1; 1 week or 2 weeks later in Experiment 2). In both experiments there was a multimedia effect (Experiment 1: $\eta_p^2 = .11$; Experiment 2: $\eta_p^2 = .15$) that was independent of the time of testing. These findings suggest that the lack of a pictorial display hindered comprehension of the pulley system's structure substantially and persistently.

1. Introduction

One of the key arguments for why it is important to investigate learning with multimedia in the first place is that a large number of studies have obtained desirable effects on learning outcomes from presenting both text and pictures (i.e., multimedia) compared with presenting text only (see Fletcher & Tobias, 2005; Levie & Lentz, 1982; Vekiri, 2002 for reviews). This well-known finding is known as "multimedia effect" (cf., Mayer, 2009) and it can be seen as a "benchmark finding" that any theory of multimedia learning needs to account for. Most studies obtained multimedia effects by assessing learning outcomes immediately after the instruction, whereas there are only a few studies testing learning outcomes after a significant delay between the instruction and testing (Bernard, Petersen, & Ally, 1981; Dwyer, 1968; Haring & Fry, 1979; Joseph, 1978; Peeck, 1974; Scheiter, Schüler, Gerjets, Huk, & Hesse, 2014). Using a delayed test to study whether knowledge has been successfully consolidated, however, is of major importance to the study of learning in general, and thus, to the study of learning with multimedia (Schweppe & Rummer, 2012).

The dearth of studies that test learning outcomes after a delay seems to mirror the role of long-term memory in theories of multimedia learning (e.g., Mayer, 2009; Schnotz & Bannert, 2003) as there is not always a clear distinction between comprehension and learning. Mostly, the focus is on the process of comprehending the materials and it is (implicitly) assumed that what is understood is what is learned and thus what is accessible when the respective information is needed at a later point in time. According to Mayer (2009, 2014), for instance, learning occurs when individuals actively engage in cognitive processing by selecting relevant information, by organizing the selected material into a coherent cognitive structure, and by integrating it with prior knowledge retrieved from long-term memory. This is exactly what is referred to as comprehension in text processing research (e.g., Graesser, Millis, & Zwaan, 1997). Learning is usually measured in terms of memory for the presented contents and transfer of the acquired knowledge, since transfer is regarded as an important indicator of successful learning (e.g., Mayer, 2009, 2014). However, learning also requires a change in long-term memory. The integrated representation of the new information and prior knowledge in long-term memory needs to be consolidated such that a durable long-term memory trace is established. When learning tests are administered immediately after acquisition, comprehension should be sufficient for good performance. With a delay between learning phase and test, comprehension is a necessary precondition for good performance, but it is no longer sufficient. In order to do well on a delayed test, it is also important not to have forgotten too much of what has been comprehended initially and to be able to retrieve this information from long-term memory. Moreover, the factors

* Corresponding author. Department of Psychology, P. O. Box 900 221, 99105 Erfurt, Germany. Tel.: +49 361 737 2203; fax: +49 361 737 2209.
E-mail address: judith.schweppe@uni-erfurt.de (J. Schweppe).

http://dx.doi.org/10.1016/j.learninstruc.2015.03.001
0959-4752/

that improve comprehension need not be those that improve retrievability. In fact, there is research demonstrating that conditions that impede performance during and immediately after learning can be beneficial for long-term learning (*desirable difficulties*, e.g., Bjork, 2013; Bjork & Bjork, 2011). As a result, assessing whether learning outcomes are not only better due to multimedia at immediate testing but remain to be so in the long-term, is important to test the viability of theories on learning with multimedia (see also Schweppe & Rummer, 2012). Moreover, assessing whether multimedia effects remain stable across time is crucial to determining the relevance of learning with multimedia for educational practice. Both in formal and in non-formal education one of the major goals is to produce sustainable knowledge gains, and not only short-lived processing advantages. Given this importance, the use of delayed tests in multimedia research is still surprisingly scarce. Accordingly, the focus of the present research is to study whether multimedia effects that may be found at immediate testing remain stable when testing after a delay (of one or two weeks).

1.1. Theoretical accounts of the multimedia effect and its stability over time

A multimedia effect is assumed to occur either because learning with text and picture results in richer and more connected memory representations than learning with text only (Baggett, 1984; Kulhavy, Lee, & Caterino, 1985; Paivio, 1991), or because the picture facilitates the comprehension process (e.g., Eitel, Scheiter, Schüler, Nyström, & Holmqvist, 2013; Schnotz & Bannert, 2003; for a similar distinction between different explanations see Schmidt-Weigand & Scheiter, 2011).

According to Baggett (1984), inspecting a picture leads to a pictorial concept that allows forming more associations with other concepts than a verbal concept, because of which concepts based on text and pictures can be retrieved more easily than concepts based on text only. *Dual coding theory* (DCT; Paivio, 1991) posits that pictures automatically yield both a non-verbal and a verbal code, whereas words only activate a verbal code. Thus, information from text is often stored in the verbal memory store only. In contrast, information from (text and) pictures will be stored in two distinct, but connected memory systems; namely a verbal and a non-verbal memory system. In this case, the same information can be retrieved either by its verbal or by its non-verbal code, and retrieval in one code automatically activates the same information stored in the other code. Therefore, retrieval of the information in only one of the two codes is sufficient for accurate recall (for a similar idea see the *conjoint retention hypothesis*, Kulhavy et al., 1985).

As the picture is assumed to benefit memory retrieval, learning with multimedia should not only improve immediate performance but should provide stable benefits. In other words, a multimedia effect should also be observable after a delay.

Alternatively, multimedia benefits can be explained such that learners are better able to *understand* an instruction when it is presented in both text and pictures than when it is presented in text only (e.g., Eitel, Scheiter, & Schüler, 2013; Eitel, Scheiter, Schüler, et al., 2013; Mayer, 2009; Schnotz & Bannert, 2003). The underlying assumption is that constructing a mental model (or situation model) represents deeper understanding (e.g., Schnotz & Bannert, 2003; Van Dijk & Kintsch, 1983). The comprehension of pictures directly results in a mental model, thereby facilitating comprehension. In contrast, constructing a mental model from text means that a propositional representation has to be constructed as an intermediate step (cf., Schnotz & Bannert, 2003). To construct a coherent mental model from text, the reader needs to infer relations that are only implicitly expressed in the text and needs to integrate text propositions with prior knowledge and/or mental images (Van Dijk & Kintsch, 1983). Thus, constructing a mental model from processing text only is effortful and error-prone. Consequently, readers may either construct a mental model that inadequately reflects the contents or situations described in a text or may even not construct a mental model at all, which is detrimental to comprehension (Schnotz & Bannert, 2003; Schnotz & Kürschner, 2008).

According to a *scaffolding assumption*, a picture can facilitate comprehension of text, because it provides more direct access to spatial information used for mental model construction (Eitel, Scheiter, & Schüler, 2013; Eitel, Scheiter, Schüler, et al., 2013). In particular, the scaffolding assumption is based on the idea that specific information about spatial relations can be directly read off a picture but have to be constructed from text by drawing inferences that are subject to interpretation (cf., Ainsworth, 2006; Schnotz & Bannert, 2003). Hence, spatial information extracted from the picture can be used as a mental scaffold for subsequent processing of (mental model construction from) text. Accordingly, when learning from text and picture, reading off information about spatial relations from the picture replaces more effortful processing of spatial information in the text. Moreover, when learning from text and picture, spatial information from the picture can constrain the range of (erroneous) interpretations of the text. Accordingly, research has shown that adding a picture to a text-based instruction fostered comprehension and sped up processing of spatial information in text (e.g., Eitel, Scheiter, & Schüler, 2013).

The accounts that attribute the multimedia effect to facilitated mental model construction do not make any specific predictions as to the stability of this benefit. Yet it can implicitly be assumed that the improved mental model construction results in a stable advantage. Moreover, due to the specificity of the picture in expressing spatial information, it can be used to resolve ambiguity that is present in text (cf., graphical constraining; Ainsworth, 2006). This reduces the likelihood of constructing an inadequate mental model from text, an advantage that should remain after a delay.

However, the construction of a mental model *during* text comprehension is a prerequisite for, but not equivalent to establishing this information as a new long-term memory trace and accessing it during delayed recall. This distinction between initial processing and comprehension on the one hand and memorability and retrievability on the other hand is particularly important in light of research in the context of *desirable difficulties* (e.g., Bjork & Bjork, 2011). This research shows that introducing difficulties in the learning phase can lead to better long-term retention (Craik & Tulving, 1975; Richland, Bjork, Finley, & Linn, 2005), sometimes at the expense of immediate performance (e.g., Roediger & Karpicke, 2006). Conversely, manipulations that facilitate performance during the encoding phase have detrimental effects on long-term retention. These findings are explained by deeper processing of the contents and a higher amount of invested mental effort triggered by the challenges that learners are faced with in the encoding phase (Bjork & Bjork, 2011; Craik & Lockhart, 1972). In line with these considerations, there are already a few studies with multimedia learning showing that effects of maximized immediate performance due to facilitated processing disappeared or even reversed once performance was assessed after a delay of one to seven days (Johnson & Mayer, 2009; Schweppe & Rummer, 2012; Segers, Verhoeven, & Hulstijn-Hendrikse, 2008; Witteman & Segers, 2010).

How can desirable difficulties be applied to the multimedia effect and its explanation via facilitated mental model construction, or, in other words, why should learning with text only function as a desirable difficulty as compared to learning with text and picture? When trying to comprehend a text without a picture, learners need