Computers in Human Behavior 60 (2016) 131–137

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Integrating written text and graphics as a desirable difficulty in long-term multimedia learning

Judith Schweppe*, Ralf Rummer

University of Erfurt, Department of Psychology, P.O. Box 900 221, 99105 Erfurt, Germany

ARTICLE INFO

Article history:
Received 29 September 2015
Received in revised form
18 January 2016
Accepted 9 February 2016
Available online xxx

Keywords:
Multimedia learning
Modality effect
Desirable difficulties
Long-term memory

ABSTRACT

Many principles for the design of multimedia learning materials share the recommendation to facilitate processing. One prominent example is the modality principle, according to which pictures should be presented with auditory rather than visual texts. Research on desirable difficulties, however, indicates that – unlike short-term learning – long-term learning benefits when processing is more demanding and therefore more effortful. In a classroom experiment (Experiment 1) and in a laboratory study (Experiment 2), we tested whether the modality principle serves long-term learning. In a multimedia presentation on the formation of lightning, we varied the text modality (oral vs. written) and the delay between learning and test (retention and transfer performance tested immediately after instruction vs. one week later). In the immediate tests, there was either an auditory advantage (Experiment 1) or no difference (Experiment 2). However, when learning was tested after a delay, the combined processing of written text and animations led to better transfer performance than oral text and animations in both experiments. This suggests that written text presentation serves as a desirable difficulty that supports long-term learning.

1. Introduction

The primary goal of research on multimedia learning is to provide recommendations for presenting learning materials. Several principles have been put forward, most of which were derived from or explained in terms of the Cognitive Theory of Multimedia Learning (CTML, Mayer, 2009, 2014) and the Cognitive Load Theory (CLT, Paas & Sweller, 2014; Plass, Moreno, & Brünken, 2010; Sweller, Ayres, & Kalyuga, 2011). Both theories focus on working memory, which is conceptualised as a limited cognitive resource, and assume that processing and learning fail whenever more capacity is needed than is available. The common core of most principles is thus to facilitate processing by reducing cognitive demands on working memory.

One influential recommendation based on this idea is to present texts in a spoken format when they accompany pictures, in order to address both modalities and therefore demand different subsystems of working memory (auditory and visual; for overviews see Ginns, 2005; Low & Sweller, 2014). According to this *modality principle*, both texts and pictures (or animations) must, at least initially, be processed in the same (visuo-spatial) working memory subsystem when texts are presented in a written format. In contrast, spoken texts and pictures (or animations) are processed in different working memory subsystems. The distribution of information among different working memory subsystems is assumed to help avoid cognitive overload (e.g., Mayer & Moreno, 1998; Moreno & Mayer, 1999). The explanation that oral and written text are stored and processed in different working memory subsystems has been criticized recently, since it is at odds with Baddeley's (e.g., 1986) working memory model and with research on working memory (Rummer, Schweppe, Fürstenberg, Scheiter, & Zindler, 2011; Rummer, Schweppe, Fürstenberg, Seufert, & Brünken, 2010; Rummer, Schweppe, Scheiter, & Gerjets, 2008; Schüler, Scheiter, Rummer, & Gerjets, 2012; Schüler, Scheiter, & van Genuchten, 2011; Tabbers, 2002). In addition (or alternatively), it is assumed that written texts and pictures/animations cannot be focussed on simultaneously and must therefore be processed sequentially. This creates a split-attention situation, which forces learners to effortfully integrate the text and picture information (Ayres & Sweller, 2014). In line with the latter explanation, the modality effect, i.e. the finding that learners perform better with a combination of spoken text and pictures than with written

* Corresponding author.
E-mail addresses: judith.schweppe@uni-erfurt.de (J. Schweppe), ralf.rummer@uni-erfurt.de (R. Rummer).

http://dx.doi.org/10.1016/j.chb.2016.02.035
0747-5632/

text and pictures, is observed less consistently with the sequential presentation of texts and pictures (e.g., Baggett & Ehrenfeucht, 1983; Tiene, 2000; but see Moreno & Mayer, 1999). A further boundary condition that applies to both explanations is the finding that the modality effect disappears (or even reverses) when learners are free to determine the pace at which they process the learning materials (e.g., Ginns, 2005; Tabbers, Martens, & van Merriënboer, 2004).

1.1. Desirable difficulties

We aim to investigate multimedia learning from a perspective that hinges on the apparent paradox of fostering learning by impeding processing rather than facilitating it and that focuses on *long-term* learning in particular. Several studies have demonstrated that cognitively more demanding learning conditions improve long-term learning, although this often comes at the expense of processing speed or initial learning (for overviews, see Pashler, Rohrer, Cepeda, & Carpenter, 2007; Roediger & Karpicke, 2006). This suggests that one cannot reliably conclude that good initial performance (as observable in an immediate test) translates into good long-term learning. One explanation for this phenomenon, which has been coined *desirable difficulties* (Bjork, 1994), is that conditions that facilitate processing induce a shallow processing mode that is particularly harmful for the establishment of stable long-term memory representations. In contrast, learning with more difficult material gives rise to processing difficulties, which consequently leads to deeper processing and representations that are less susceptible to forgetting (Bjork, 2013; Craik & Lockhart, 1972).

There is one example concerning the presentation of materials in multimedia research that can be regarded as a desirable difficulty: De Croock, van Merriënboer, and Paas (1998) presented learners with practice problems in a random schedule with problems that varied with respect to the procedures necessary for performing each task compared to a blocked schedule. The random schedule *increased* the complexity and the cognitive demands but resulted in better transfer performance.

While in tests administered immediately after instruction short-lived advantages of the easier conditions often outweigh the beneficial effects of difficult encoding, those latter effects are particularly strong when it comes to long-term tests of learning (Pashler et al., 2007; Roediger & Karpicke, 2006). Given that several recommendations for the design of multimedia learning materials emphasize the facilitation of processing, this may imply that what is beneficial for short-term multimedia learning sometimes hinders long-term learning. In research on multimedia learning, the literature is quite clear that an increase in processing demands usually impedes rather than aids learning. Yet maybe the negative effects that cognitive load has on initial understanding only mask its positive effects when learning outcomes are restricted to a condition of immediate testing. Effects attributed to desirable difficulties are stronger and often only present when performance is tested after a delay, that is, after forgetting could have occurred (e.g., Roediger & Karpicke, 2006). If this is the case, conditions that are beneficial in immediate tests of multimedia learning might have adverse effects in the long term and the respective principles may need to be revised for long-term learning.

One condition for which this could be the case is the processing of written versus oral text in combination with pictures. A situation in which learning materials must be processed sequentially – which is the case with visual texts plus pictures/animations – forces the learner to switch between text and picture. It is thus necessary to retain the pictorial information while reading the text and the verbal information while processing the picture (Rummer et al., 2008, 2010). This causes learners to effortfully integrate the information sources – a process that learners are likely to engage in when they are motivated and that may underly the higher cognitive load associated with written compared to oral texts (e.g., Ayres & Sweller, 2014). Such a demanding integration process may be harmful for initial processing, as evident by the immediate modality effect, but may have the advantage of leading to deeper processing in terms of the levels-of-processing hypothesis (Craik & Lockhart, 1972), and therefore be particularly important for long-term memory. By presenting written text along with pictures or animations, learners might be "tricked" into deeper processing, which, in turn, results in more stable long-term memory representations. Of course, such positive long-term benefits of difficult learning conditions can only occur in cases in which the difficulty impedes processing but does not entirely preclude comprehension.

1.2. Long-term tests of multimedia learning

In surveying the literature on multimedia learning, we found only few experiments in which learning was measured after a noticeable delay. With respect to the modality effect that we investigate in our study, we found four such studies (Segers, Verhoeven, & Hulstijn-Hendrikse, 2008; She & Chen, 2009; Van den Broek, Segers, & Verhoeven, 2014; Witteman & Segers, 2010). Segers et al. (2008) observed an inverse modality effect (i.e., an advantage for written text) in a transfer test one week after the learning phase. Witteman and Segers (2010) observed a written text advantage for a transfer test after one day and no modality differences one week later. She and Chen (2009) observed a written advantage for retention after a five-week delay between encoding and test and Van den Broek et al. (2014) after a delay as short as one day. These findings are important first steps towards a test of the long-term consequences of reading versus listening. Remarkably, none of the delayed tests revealed a modality effect in terms of an advantage for oral texts – there was either no modality difference or a written advantage. However, these inverse modality effects do not suffice to question the applicability of the modality principle to long-term learning since in all studies the presentation of the materials was learner-paced. As indicated above, learner-paced presentation often results in visual advantages even with immediate tests and is therefore a boundary condition of the modality effect (e.g., Ginns, 2005; Tabbers et al., 2004). Consequently, further research is imperative to determine whether reading is advantageous even with system-paced presentation.

1.3. Objectives of the present study

We aim to investigate the modality effect after a delay, but under conditions in which it is otherwise highly likely to occur. That is, in contrast to previous studies (e.g., Segers et al., 2008), we will test the modality principle with experimenter-paced learning. We will compare participants' performance on retention and transfer tests after multimedia lessons with auditory versus visual text (based on the lightning materials by Mayer & Moreno, 1998; Moreno & Mayer, 1999) when these tests are administered either immediately or one week after the instruction. According to the modality principle, there should be an advantage of auditory over visual texts even with a delayed test. According to the desirable difficulties approach, however, visual text presentation should lead to superior performance on the delayed test because learners were forced to process the texts more effortfully and more deeply, leading to more stable memory representations that are less susceptible to forgetting.