Multimedia in the wild: Testing the validity of multimedia learning principles in an art exhibition

Stephan Schwan[*], Silke Dutz, Felix Dreger

Leibniz-Institut fuer Wissensmedien, Schleichstr. 6, 72076 Tuebingen, Germany

1. Introduction

Multimedia learning and knowledge acquisition not only takes place in formal education but is also found outside of schools and universities in a wide range of non-formal and informal situations and settings. The spectrum of informal occasions of multimedia learning ranges from reading an illustrated popular science journal or watching a documentary on TV to exploring a multimedia Internet site related to one's favorite hobby or using an electronic guide during a museum visit. Although the basic cognitive mechanisms of processing multimedia content can be assumed to operate in similar ways across different situations, informal and formal learning circumstances differ in characteristics that may influence the manner of how knowledge acquisition takes place. More specifically, informal learning is usually intrinsically motivated and proceeds voluntarily on a free-choice basis. Therefore, learners' interest in the material may quickly vanish, but their attention may also hold over an extended period of time, leading to a broader inter-individual distribution of processing time than normally found under controlled laboratory conditions. Also, due to the free-choice character of informal settings, knowledge acquisition typically proceeds in a self-determined manner, which has been found to reduce cognitive effects of principles of multimedia learning (Ginns, 2005; Tabbers, Martens, & van Merrienboer, 2004). Finally, on informal occasions, learners' often not only focus on knowledge acquisition but simultaneously pursue other motives as well, such as making an aesthetic experience or being entertained (Glaser, Garsoffky, & Schwan, 2012; Pekarik, Doering, & Karns, 1999). Hence, conflicts between the instructional and the aesthetical or entertaining value of multimedia material may arise.

The present study was conducted to address these issues by investigating the effect of various combinations of text with static pictures, that is, multimedia material, on visitors' behavior and knowledge acquisition in an art exhibition setting. Visitor studies have shown that learning and understanding are important motives for visiting museums and exhibitions (Falk, 2009; Pekarik et al., 1999). Therefore, art exhibitions may be seen as informal learning settings in which visitors acquire some knowledge about the exhibit's topic (accordingly, we will use the terms visitor and learner interchangeably throughout the text).

We investigated whether established principles and findings from laboratory multimedia research on the effects of combining pictures with written or spoken text on learning, namely, the multimedia principle and the modality principle, can also be found in such an informal learning setting. In addition, we tested whether the way of providing text information (permanent presentation versus active access) has an effect on learning. Finally, we also tested whether these principles have an effect on the average time spent with the artworks, whether average time spent with the artworks had an influence on retention, and whether the principles interfere with the visitors' aesthetic appreciation of the artworks.

2. Theoretical background

2.1. Principles of multimedia learning

Principles of multimedia learning have been developed on the basis of Cognitive Theory of Multimedia Learning (Mayer, 2009) and its extension to Cognitive Affective Theory of Multimedia Learning (CATML, Moreno & Mayer, 2007) as well as of Cognitive Load Theory (Sweller, Ayres, & Kalyuga, 2011). According to these theories, learning is considered a multi-step process that requires elaborate processing of the content to be learned in working memory, which is resource-limited. Working memory is considered to be partitioned into several substructures, each with its own limited processing resources (Baddeley, 2012). Two of these working memory substructures are devoted to the separate processing of visual and auditive input, each separately. In working memory, the input of both channels is integrated and subsequently transferred into permanent knowledge structures in long-term memory.

In its general formulation, the *multimedia principle* proposes a learning advantage of presenting content via a simultaneous combination of pictorial and textual or verbal material over presenting content via one type of presentation alone (Butcher, 2014; Mayer, 2009). More specifically, whereas in its initial formulation a learning advantage of text plus picture over text alone was postulated, recently it has been extended to a learning advantage of picture plus text over picture alone (Glaser & Schwan, 2015). At least two cognitive mechanisms contribute

to this advantage. First, the processing of both pictorial and textual information proceeds in separate substructures in working memory, which after integration lead to a more elaborated mental representation of the content. Also, Glaser and Schwan (2015) have argued that in the case of pictorial information, accompanying text information may additionally define an interpretive frame for the picture and may also guide viewers' attention to relevant parts of the picture, again fostering comprehension in comparison to a picture only condition. Accordingly, the validity of the multimedia principle has been demonstrated in numerous laboratory studies and is thus one of the best-established principles in the design of instructional material (Butcher, 2014).

The *modality principle* is one important specification of the multimedia principle. In its general formulation, it proposes a learning advantage of accompanying pictorial material by spoken instead of written text for two reasons (Low & Sweller, 2014). First, the presentation of content to be learned via two sensory channels leads to a more balanced use of the working memory's resources (Kalyuga, Chandler, & Sweller, 1999). Second, the necessity of switching between spoken text and picture is minimized because auditive and visual information can be attended to in a simultaneous manner (Tabbers et al., 2004). Yet, empirical findings have indicated that this principle primarily holds for the provision of short texts under system-controlled conditions (Ginns, 2005; Tabbers et al., 2004), whereas for lengthy texts, written texts have the advantage of providing a higher flexibility of use (for example, in terms of pace or re-readings) and can therefore be better adapted to the cognitive processing needs of the learner (Leahy & Sweller, 2011; Tabbers et al., 2004). In contrast, accompanying text has also a strong attention guiding function (Glaser & Schwan, 2015), which is presumably better served by a spoken than by a written text because the latter may suffer due to the problem of split attention. Therefore, at least for complex static visual depictions, the spoken text's advantage of guiding viewers' attention may outweigh the disadvantages of a lower flexibility of use.

Finally, Moreno and Mayer (2007) have outlined an extended model of multimedia learning (CATML) that not only includes metacognition, motivation, and affect as determinants of multimedia learning but also introduces issues of design of interactive features in multimodal learning environments. They distinguish between different types of interactivity, ranging from dialoguing (e.g., receiving questions and answers), to controlling (e.g., determining pace and order), manipulating (e.g., setting parameters) and navigating (moving between content areas) to searching (accessing new content material). In the present context, the searching type of interactivity is of particular importance because multimedia presentations in museums may either present additional written text information permanently beside the exhibits or require visitors to actively search and access the information via a digital mobile guide. According to CATML, having learners to actively access the text information on a digital guide instead of providing this information permanently should impose extraneous processing, which is defined as "cognitive processes that are not necessary for making sense of the new information but instead originated from poorly designing the task" (Moreno & Mayer, 2007, p. 314). The concept of extraneous processing bears a strong resemblance with the notion of extraneous cognitive load that has been formulated in Cognitive Load Theory (Sweller et al., 2011). According to Cognitive Load Theory, suboptimal information design may place unnecessary additional demands on working memory, thereby reducing memory resources which are required for information elaboration and learning. Hence, the use of an interactive digital guide instead of permanent labels should reduce the cognitive resources available for processing the learning content.

2.2. Multimedia learning in informal settings

Principles of multimedia learning have been investigated and validated in a large number of studies, both in controlled laboratory settings and under field conditions in classroom settings (Tabbers et al., 2004; Harskamp, Mayer, & Suhre, 2007; Issa et al., 2013; Scheiter, Schüler, Gerjets, Huk, & Hesse, 2014). In contrast, empirical investigations of the role of multimedia principles in informal learning settings is to our knowledge largely absent. Yet, due to their presumed generality, principles of multimedia learning should also apply to learning situations outside of the classroom. In particular, museums and exhibitions have been considered typical examples of informal learning settings (Bell, Lewenstein, Shouse, & Feder, 2009; Hodkinson, Colley, & Malcolm, 2003). Museums are important institutions for providing knowledge not only during school field trips or family visits but also as an opportunity for informally acquiring knowledge about science and technology as well as about history, art, and culture in adulthood (Falk, Storksdieck, & Dierking, 2007; Mastandrea, Bartoli, & Bove, 2007; Schwan, Grajal, & Lewalter, 2014).

Informal learning settings, such as museums and exhibitions, are not only characterized by taking place outside of formal educational institutions but differ from classroom settings on a number of points (Bell et al., 2009; Schwan et al., 2014). In particular, the contents of a museum exhibition are usually not part of a specific curriculum; the visitors come to the exhibition without specific learning goals, and they do not receive any grading. Thus, knowledge acquisition in museums is usually intrinsically motivated either by personal interest or by situational interest that spontaneously develops during the visit, being manifested in visitors' self-determined selection of the exhibits that they choose to give attention to. Accordingly, knowledge acquisition in museums has been termed "free-choice" learning (Falk & Dierking, 2000), which implies that both the selection of exhibits and how deeply they are explored and elaborated is left to the visitor.

One important implication of conditions of "free-choice" learning is that different types of multimedia presentations may not only shape *how*, but also *how long*, pictorial artworks are processed by the learners. For example, in several museums increased dwell times in exhibitions for visitors using audio guides or digital guides have been shown (for audio guides: Webb & Mann, 2014; for digital mobile guides: Eghbal-Azar, Merkt, Bahnmueller, & Schwan, 2016; Lanir, Kuflik, Dim, Wecker, & Stock, 2013). While in formal learning contexts learning time is interpreted as an indicator of information elaboration (e.g. Spanjers, van Gog, & van Merrienboer, 2010), visitor research usually considers dwell time as a proxy for interest and motivation (Boisvert & Slez, 1995; Sandifer, 2003; Serrell, 1997). However, both interpretations should not be seen as mutually exclusive, but as complementary instead. Accordingly, visitors reported that the additional information provided by a digital mobile guide (Helal, Maxson, & Ancelet, 2013) or audio guide (Webb & Mann, 2014) helped them "to slow down" (indicating increased interest) and supported them to take a more focused look at an artwork (indicating increased elaboration). Taken together, these findings suggest that in free-choice learning contexts, time spent looking at an exhibit should not be simply considered a methodological confound, but instead might constitute an independent effect of multimedia design principles on learning processes.

Given these specifics of learning in informal contexts, demonstrating the validity of design principles derived from theories of multimedia learning for museum settings would substantially enlarge the boundary conditions under which these principles may be applied.

2.3. Using media to explain exhibits in museums

In order to describe and explain their exhibits, museums use various types of media, including written texts (for example, via interpretive labels), spoken texts (for example, via audio guides), and also increasingly digital guides (Proctor, 2011; Tallon & Walker, 2008). The spectrum of media providing additional information for the visitors is particularly broad in art museums and exhibitions. Some art museums restrict additional information to small labels which only specify title, artist, and the year of the painting, arguing that extended explanations