Learning and Instruction 58 (2018) 274–283

Contents lists available at ScienceDirect

Learning and Instruction

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

Effectiveness of eye movement modeling examples in problem solving: The role of verbal ambiguity and prior knowledge

Tim van Marlen[a,*], Margot van Wermeskerken[a], Halszka Jarodzka[b,c], Tamara van Gog[a]

[a] Department of Education, Utrecht University, The Netherlands
[b] Welten Institute, Open University of the Netherlands, The Netherlands
[c] Humanities Laboratory, Lund University, Sweden

ARTICLE INFO

Keywords:
Example-based learning
Eye tracking
Attention cueing
Eye movement modeling examples

ABSTRACT

Eye movement modeling examples (EMME) are video modeling examples with the model's eye movements superimposed. Thus far, EMME on problem-solving tasks seem to be effective for guiding students' attention, but this does not translate into higher learning outcomes. We therefore investigated the role of ambiguity of the verbal explanation and prior knowledge in the effectiveness of EMME on geometry problems. In Experiment 1, 57 university students observed EMME or regular video modeling examples (ME) with ambiguous verbal explanations. Eye-tracking data revealed that –as in prior research with unambiguous explanations- EMME successfully guided students' attention but did not improve test performance, possibly due to students' high prior knowledge. Therefore, Experiment 2, was conducted with 108 secondary education students who had less prior knowledge, using a 2 (EMME/ME) x 2 (ambiguous/unambiguous explanations) between-subjects design. Verbal ambiguity did not affect learning, but students in the EMME conditions outperformed those in the ME conditions.

1. Introduction

Video modeling examples in which a model demonstrates and explains how to perform a task (e.g., "how to" tutorial videos on YouTube), are widely used in formal and informal learning settings. Such videos lie in the tradition of example-based learning, which is an effective and efficient way of learning, provided that the examples are well-designed (Renkl, 2014; Van Gog & Rummel, 2010). It has been proposed that, depending on the task at hand, the design of screen-recording video examples could be further improved by showing learners what the model is looking at, by displaying the model's eye movements (Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). Displaying a visualization of the model's eye movements (e.g., fixations represented as a circle or dot) is expected to guide learners' attention to what the model is looking at in that moment, which should make it easier to understand and learn from the demonstration and verbal explanation. Several studies have found beneficial effects of such "eye movement modeling examples" (EMME) on attention guidance and found enhanced learning of classification tasks (Jarodzka et al., 2012; Jarodzka, Van Gog, Dorr, Scheiter, & Gerjets, 2013) and enhanced integration of text and pictures during reading (Mason, Pluchino, & Tornatora, 2015a; 2015b). Thus far, however, EMME on problem-solving tasks seem to be effective for guiding learners' attention, but this does not translate into higher learning outcomes (i.e., higher performance on the test problems) compared to the no EMME control condition (Van Marlen, Van Wermeskerken, Jarodzka, & Van Gog, 2016). This discrepancy between the studies regarding the effectiveness of EMME might be related to the extent to which the verbal explanation accompanying the EMME is clear (i.e., unambiguous) to the participants. For instance, studies using classification tasks in which EMME were found to be effective (Jarodzka et al., 2012, 2013), specific jargon was used; when learners do not yet know the jargon, this increases the usefulness of visual guidance. In contrast, in a study on learning to solve a puzzle problem, it was clear from the verbal explanation what object the model was looking at, and the visual guidance provided by EMME was not useful for learning (Van Gog et al., 2009). The present study addressed two potential explanations for this lack of effect of attention guidance on learning procedural problem-solving from EMME: ambiguity of the verbal explanation and prior knowledge.

1.1. Eye movement modeling examples

Multimedia materials provide a combination of verbal and pictorial information, which according to the *dual-coding theory* (Clark & Paivio, 1991; Paivio, 1986) are processed in separate auditory and visual channels. According to the *Cognitive Theory of Multimedia Learning*

* Corresponding author. Department of Education, Utrecht University, P.O. Box 80140, 3508 TC, Utrecht, The Netherlands.
E-mail address: t.v.a.vanmarlen@uu.nl (T. van Marlen).

https://doi.org/10.1016/j.learninstruc.2018.07.005
Received 6 November 2017; Received in revised form 12 July 2018; Accepted 15 July 2018
0959-4752/

(Mayer, 2014), learners first need to attend the relevant verbal and pictorial information (selection). After selecting the relevant verbal and pictorial information learners *organize* this information into coherent mental representations, and *integrate* the verbal and pictorial mental representations with each other and with available prior knowledge (Mayer, 2014). In dynamic learning materials like video modeling examples, one challenge for the selection of information lies in the transience of the material. If the learner does not attend to the right information at the right moment, it is no longer available for processing (i.e., organization and integration) and learning is hindered (Ayres & Paas, 2007). One reason why learners might not be able to attend to the right information at the right time in a video modeling example, is that it is likely that there is a discrepancy in what the expert model and the novice learner are attending to, and that the verbal explanation provided by the model may not be sufficiently clear to rapidly guide the learner's attention to what the expert is referring to.

The discrepancy between experts' and novices' attention allocation has been shown in different eye-tracking studies. Experts often attend to task-relevant information relatively longer and faster while paying less attention to task-irrelevant information than novices (Charness, Reingold, Pomplun, & Stampe, 2001; Van Gog, Paas, & Van Merriënboer, 2005; Van Meeuwen et al., 2014; Wolff, Jarodzka, Van den Bogert, & Boshuizen, 2016). This expertise effect has also been demonstrated within participants as a result of task experience (Blair, Watson, & Meier, 2009; Canham & Hegarty, 2010; Haider & Frensch, 1999; Hegarty, Canham, & Fabrikant, 2010). This difference in attention allocation might cause learners to miss the information the model is attending to, unless the model's verbal explanation would be sufficiently clear to rapidly guide learners' attention to the right information at the right time.

It has been proposed that one way to reduce the discrepancy between the model's and the learner's attention allocation would be to show the learner what the model is attending to, by displaying a visualization of the model's eye movements (e.g., as a dot or circle; Van Gog et al., 2009). In such eye movement modeling examples (EMME), the learner is not only shown how the model is performing the task (by means of a screen recording of the model's computer screen), but also where the model was looking while performing the task. By showing the eye movements of the model the visual attention of the learners is guided and synchronized with the model to create a state of *joint attention* (i.e., *joint attention* is the phenomenon characterized as automatically looking at an object someone else is looking at; Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008; Frischen, Bayliss, & Tipper, 2007), thus helping the learners attend to the relevant information at the right time which, in turn, can be expected to improve learning.

However, findings regarding the effectiveness of EMME for learning are mixed. Whereas some studies have found beneficial effects on learning classification tasks (Jarodzka et al., 2012, 2013) and learning text and picture integration (Mason et al., 2015b, 2015a), EMME in which problem-solving tasks are demonstrated seem to be effective for guiding learners' attention (Van Marlen, Van Wermeskerken, Jarodzka et al., 2016), but this does not translate into higher learning outcomes (Van Marlen, Van Wermeskerken, Jarodzka et al., 2016; see also; Van Gog et al., 2009). One possible reason for these mixed findings might lie in the extent to which the model's verbal explanation is sufficient to rapidly guide learners' attention to what the model is referring to. EMME might be most effective for learning when the model's verbalizations are ambiguous.

1.2. The role of verbal ambiguity in the effectiveness of EMME

When the verbal explanation in a modeling example contains ambiguous verbal referents, it will not be immediately clear to the learner what the model is referring to. Ambiguity of verbal referents can originate from different sources. For instance, experiments in cognitive science have shown that ambiguity can arise due to the visual context (e.g., multiple objects that the referent might refer to; Allopenna, Magnuson, & Tanenhaus, 1998; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995), the lack of specificity of the verbal information (e.g., referring to a target without verbal location descriptions; Louwerse & Bangerter, 2010), or a lack of prior knowledge (e.g., lack of background knowledge about a to be inspected and discussed painting; Richardson, Dale, & Kirkham, 2007; Experiment 2). When verbal referents are ambiguous for any of those reasons, it will take listeners more time to locate the relevant (i.e., target) information, if they are able to locate it at all (Louwerse & Bangerter, 2010; Van Marlen, Van Wermeskerken, & Van Gog, 2018).

These studies about the effects of verbal ambiguity on the speed and accuracy with which referents are located, suggest that the attention guidance provided by EMME might be most needed and most effective for learning when the model's verbal explanation is ambiguous for learners. Providing clear verbal explanations might not always be possible depending on the task and source of the ambiguity. For instance, in classification tasks providing a clear verbal description of a complex visual shape denoted by a jargon term, might be quite difficult. In the classroom, teachers/instructors can resolve this problem by using available non-verbal cues (e.g., looking or pointing at the part of the task they are discussing) that will disambiguate their verbal message. However, in digital video instructions these non-verbal cues are not necessarily present. Thus, it is likely that verbal explanations in some circumstances are not sufficient and have to be accompanied by non-verbal cues that align the learners' attention with that of the model. EMME do this by showing the learner what the model is looking at, at any given moment, which may resolve potential ambiguities in the model's verbal explanation. Hence, the discrepancy in results regarding the effectiveness of EMME might be due to whether verbal explanations are perceived as ambiguous without further guidance of an EMME. Indeed, there is some tentative evidence suggesting that this is the case: It is likely that verbal referents were ambiguous for the learners in the studies that found positive effects of EMME on learning classification tasks (Jarodzka et al., 2012, 2013; Vitak, Ingram, Duchowski, Ellis, & Gramopadhye, 2012).

For instance, in one study learners had to learn to classify fish locomotion patterns and were shown videos of different fish while an expert gave verbal explanations about their locomotion pattern (Jarodzka et al., 2013). In his verbal explanation, the expert was using terms like 'the dorsolateral fin', which can be ambiguous when the learner does not yet know what that is. In this case, seeing the expert's eye movements (i.e., seeing what he is looking at) would help to attend to the right information at the right time. Similarly, in the study by Vitak et al. (2012) learners had to classify cells on histological slides. The learners were shown video examples with or without the expert model's eye movements superimposed onto the example slides while listening to a verbal explanation. The expert referred to certain cells with terms like "there's" or "this one". Results indicated that the expert's eye movements were helpful in disambiguating the verbal referents as indicated by fewer classification errors on subsequent test tasks. Also, search behavior was more efficient for learners in the EMME condition as those learners needed less time to classify the cells and made fewer fixations on subsequent test tasks.

In studies on problem-solving tasks, in which the verbal referents were likely unambiguous (e.g., referring to "angle A" of a geometry problem), EMME had no beneficial effects on learning (Van Marlen, Van Wermeskerken, Jarodzka et al., 2016) or a negative effect (Van Gog et al., 2009). In the study by Van Marlen, Van Wermeskerken, Jarodzka et al. (2016), participants were shown modeling examples with or without the model's eye movements superimposed while also hearing verbal explanations about how to solve geometry problems (Exp. 2). Although EMME were effective for *more rapidly* guiding attention towards the information the model referred to (i.e., shorter times to first fixations and –probably as a consequence– longer fixation of referents), there was no difference in how many referents were fixated between the