ELSEVIER

Learning and Instruction 20 (2010) 167–171

Learning and Instruction

www.elsevier.com/locate/learninstruc

Commentary

Unique contributions of eye-tracking research to the study of learning with graphics

Richard E. Mayer*

Department of Psychology, University of California, Santa Barbara, CA 93106-9660, USA

Abstract

The author examines the empirical, methodological, theoretical, and practical contributions of the six studies in this special issue on eye tracking as a tool to study and enhance multimedia learning. The design of learning environments involving graphics should be consistent with a research-based theory of how people learn and evidence-based principles of how to help people learn. Research using eye tracking offers a unique path to testing aspects of theories of multimedia learning, particularly concerning perceptual processing during learning. The studies reported in this special issue add to the evidence base on how people learn and think with graphics.

Keywords: Multimedia learning; Eye-tracking; Computer-based instruction; Graphics comprehension

1. Introduction

An essential aspect of research on multimedia learning (i.e., learning with words and graphics) concerns how people learn and think with graphics (i.e., the science of learning) and how to help people learn and think with graphics (i.e., the science of instruction). These issues are addressed in this special issue.

Research on instructional methods can address three kinds of questions: what works, when it works, and how it works. Concerning what works, recent advances in the science of instruction have produced a set of evidence-based principles for how to design lessons containing words and pictures (Clark & Mayer, 2008; Mayer, 2005, 2009; O'Neil, 2005, 2008). For example, the signaling principle is that people learn more deeply from a multimedia lesson when essential material is highlighted and the modality principle is that people learn more deeply from animation and narration than from animation and onscreen text (Mayer, 2009). Research reported in this special issue contributes to this effort to determine what works by examining the effectiveness of four instructional techniques (such as signaling or modality) on measures of learning outcome or cognitive performance.

Concerning when it works, there is increasing evidence pinpointing the boundary conditions for instructional-design principles, that is, pinpointing for whom, for which kinds of materials, and under which learning situations each principle applies. Research reported in this special issue contributes to this effort by distinguishing boundary conditions for the modality principle and the signaling principle.

Concerning how it works, the goal is to describe the learning process that underlies the effectiveness of an instructional method. The eye-tracking methodology used in each of the studies in this issue offers a unique opportunity to contribute to understanding the learner's perceptual processing during learning. Thus, in this issue the authors go beyond asking simply "what works?" or "when does it work?" The major new contribution of the work reported here concerns using eye-tracking methodology to address our third goal, that is, to determine how a particular instructional method causes learning. In this way, the research in this special issue contributes both to the science of instruction by offering new ways of determining what works, and the science of learning by helping test implications of learning theory concerning perceptual processing during learning.

* Corresponding author. Tel.: +1 805 8932472; fax: +1 805 8934303.
E-mail address: mayer@psych.ucsb.edu

0959-4752/$ - see front matter
doi:10.1016/j.learninstruc.2009.02.012

2. Empirical contributions

Table 1 lists some of the main features of the experimental comparisons reported in this issue. In particular, I focus on independent variables that have direct instructional design implications, that is, signaled vs. non-signaled graphics (signaling), lower vs. higher knowledge learners (prior knowledge), animation-and-narration vs. animation-and-onscreen text (modality), and fast-to-slow vs. slow-to-fast pacing (pacing). I also focus on dependent measures involving eye-fixations, namely total fixation time on relevant areas of a graphic, and dependent measures involving learning outcome or cognitive performance such as transfer test score. As you can see all of the studies involve graphics – including animation, video, and static illustrations – but only one involves accompanying words (Schmidt-Weigand, Kohnert, & Glowalla, 2010).

Do the independent variables have effects on the two kinds of dependent measures? Table 2 summarizes the answer to this question for each of the six projects.

The first two studies (by Boucheix & Lowe, 2010, and by De Koning, Tabbers, Rikers, & Paas, 2010, respectively) found a *signaling effect* for measures of eye-fixation time and measures of learning (or deep cognitive processing during learning): students who viewed an appropriately signaled animation spent more time viewing the relevant areas of an instructional animation and scored higher on measures of mental model construction than did students who viewed a non-signaled animation. Although Boucheix and Lowe found significant differences between signaled and non-signaled treatments on learning measures of overall amount learned, De Koning et al., 2010.

The second two studies (by Jarodzka, Scheiter, Gerjets, & Van Gog, 2010, and by Canham & Hegarty, 2010) found a *prior knowledge effect* both for measures of eye-fixation time and measures of cognitive performance: high knowledge students spent more time viewing the relevant areas of an instructional graphic and scored higher on accuracy in answering questions that required inferences based on the graphic.

The fifth study (by Schmidt-Weigand et al., 2010) found consistent evidence for a *modality effect* for measures of eye-fixation time but not for measures of learning outcome such as a transfer test. How can we account for the lack of modality effect on the transfer test? In Experiment 1, the authors chose to change the transfer test from individually-timed open-ended questions (which had been successfully used in numerous previous experiments) to multiple-choice questions (which had not be validated in previous studies) and thereby may have diminished the sensitivity of the primary measure of learning outcome. In Experiment 2, learners were given control over the pace of presentation, which has been shown in previous research to diminish the modality effect (Mayer, 2009).

The sixth study (by Meyer, Rasch, & Schnotz, 2010) did not find consistent evidence for either version of a *pacing effect* – presenting an animation at a fast rate followed by a slow rate results in better (or worse) overall learning than presenting an animation at a slow rate followed by a fast rate – either on eye-fixation or on learning outcome measures. However, the authors report an interesting interaction on the comprehension test in which fast-to-slow pacing results in relatively smaller differences between micro-knowledge and macro-knowledge compared to slow-to-fast pacing.

In summary, the authors of papers in the present special issue have been able to show that the signaling effect, prior knowledge effect, and to some extent, the modality effect can be extended to measures of eye-fixation time.

3. Methodological contributions

To conduct useful research on instructional effectiveness, it is necessary to have clearly defined independent variables (i.e., instructional methods) and clearly specified dependent

Table 1
Content, medium, main independent variable, main eye-fixation measure, and main learning outcome measures for the six eye-tracking studies in the present special issue.

Study	Content	Medium	Main independent variable	Eye fixation measure	Main learning outcome measures
Boucheix and Lowe, 2010	Piano	Animation only	Signaling: spreading colors vs. arrows vs. no arrows	Time looking at relevant areas	Comprehension score (including configuration, local kinematics, and mental model score)
De Koning et al., 2010	Heart	Animation only	Signaling: shading cues vs. no cues	Time looking at relevant areas	Comprehension and transfer scores; number of explanatory statements in verbal protocols
Jarodzka et al., 2010	Swimming fish	Video only	Expertise: professors vs. students	Time looking at relevant areas	Accuracy in describing locomotion
Canham and Hegarty, 2010	Weather maps	Static map only	Prior knowledge: before vs. after instruction	Time looking at relevant areas	Accuracy in verifying wind direction
Schmidt-Weigand et al., 2010	Lightning	Animation and words	Modality: animation with narration vs. animation with onscreen text	Time looking at relevant areas	Comprehension, transfer, visual memory scores
Meyer et al., 2010	Four-stroke engine	Animation only	Presentation rate: fast vs. slow	Time looking at relevant areas	Comprehension score

Table 2
Main empirical and theoretical contributions of the six eye-tracking studies in the present special issue.

Study	Main empirical contribution	Main theoretical implication
Boucheix and Lowe, 2010	Learners spend more time looking at relevant areas of an animation when relevant features are highlighted both spatially and temporally (signaling effect).	Visual signals guide the learner's visual attention.
	Signaled group performs better on comprehension test involving low-salience/high-relevance features (signaling effect).	Strong link between eye-fixations and learning outcomes.
De Koning et al., 2010	Learners spend more time looking at relevant areas of an animation when the relevant features are highlighted (signaling effect).	Visual signals guide the learner's visual attention.
	No differences on retention or transfer scores effects (no signaling effect); signaled group produces more explanatory statements on verbal protocol (signaling effect).	Inconsistent link between eye-fixations and learning outcomes: dependent measures may not be sensitive.
Jarodzka et al., 2010	Experts spend more time looking at relevant areas of video than do novices.	Prior knowledge guides visual attention.
	Experts are more accurate in describing the fish's locomotion than are novices.	Strong link between eye-fixations and learning outcomes.
Canham and Hegarty, 2010	Learners spend more time looking at relevant areas after instruction.	Prior knowledge guides visual attention.
	Learners are more accurate in determining wind direction after instruction.	Strong link between eye-fixations and learning outcomes.
Schmidt-Weigand et al., 2010	Learners spend more time looking at relevant areas of an animation when they receive animation and narration rather than animation and onscreen text (modality effect).	Learners who receive animation and onscreen text must split their attention between graphics and printed words.
	No differences on retention or transfer scores (no modality effect); animation and narration performs better on visual memory test when pacing is slow (modality effect).	No strong link between eye-fixations and learning outcomes: slow or learner-paced presentation rate allowed learners enough time in all treatments; dependent measures may not be sensitive.
Meyer et al., 2010	No strong effects of presentation pace (fast-to-slow vs. slow-to-fast) on eye fixations.	No strong evidence that fast system-imposed presentation rate primes attention to macro-events and slow system-imposed presentation rate primes attention to micro-events.
	No strong effects of presentation pace on overall comprehension; interaction showing relatively less difference between micro- and macro-knowledge for fast-to-slow pacing than slow-to-fast pacing.	Mixed evidence that presentation pace affects what is learned.

measures (i.e., measures of perceptual processing during learning, cognitive processing during learning, and learning outcomes). Concerning independent variables, the studies reported in this issue clearly specify modality (i.e., animation and narration vs. animation and onscreen text) and pacing (i.e., fast presentation rate followed by slow presentation vs. slow presentation rate followed by fast presentation rate), help to sharpen the critical features of visual signaling that are effective (i.e., the use of spreading colored areas and the use of shading out of irrelevant areas), and exemplify alternative ways of manipulating prior knowledge (i.e., both as a between subjects factor and a within subjects factor).

Concerning dependent measures, the single most important methodological contribution of the studies in this issue is to show that eye-tracking measures, such as total fixation time on relevant areas of an instructional graphic, can be successfully added to researchers' toolboxes as a way of testing hypotheses about perceptual processing during learning under different instructional methods. In short, the authors demonstrate a useful methodology for measuring perceptual processing during learning and thinking with graphics.

However, it is clear that a major challenge of conducting useful research on instructional methods is to develop appropriate measures of learning outcome and cognitive processing during learning. For example, one technique for evaluating cognitive processing during learning is to use a subjective questionnaire to assess the learner's cognitive load (De Koning et al., 2010). However, recent research has shown that different ways of measuring cognitive load may be tapping different kinds of cognitive processing during learning (DeLeeuw & Mayer, 2008), so the search for valid and reliable measures of cognitive load continues. Another technique for measuring cognitive processing during learning is using cued retrospective reporting in which learners describe what they were thinking as they watch a video showing their sequence of eye fixations (De Koning et al., 2010; Jarodzka et al. 2010). However, developing a scoring rubric becomes a major challenge; for example, de Koning and colleagues focus on statements about relevant subsystems such as valves in the heart, whereas Jarodzka and colleagues focus on relevant underlying concepts such as statements about fish species. It is clear that more work is needed on developing a consistent approach to categorizing learners' statements on protocols.

Finally, in order to measure learning outcomes some projects used multiple-choice comprehension tests (e.g., De Koning et al., 2010; Meyer et al., 2010; Schmidt-Weigand et al., 2010), but these multiple-choice tests generally were not sensitive to differences in instructional treatments. Thus, there may be justification for using short open-ended questions that have a fixed answer time and are scored by tallying the number of acceptable answers, which have been sensitive in prior multimedia research (Mayer, 2009). The most successful measure of learning outcome was Boucheix and Lowe's (2010) mental model comprehension score which involves writing an answer to an open-ended question and scoring it based on strict rubric. Determining how to assess what someone knows remains a central challenge of instructional research (Anderson et al., 2001; Pelligrino, Chudowski, & Glaser, 2001).

4. Theoretical contributions

A serious challenge for eye-tracking researchers is to find the sometimes-missing link between eye-fixation measures and learning outcome (or cognitive performance) measures. In the present set of six studies, total fixation time on relevant areas of a graphic is intended as a measure of perceptual processing, and is hypothesized to cause cognitive processing that leads to better learning (or cognitive performance). Table 3 summarizes whether there was a link between eye-fixation measures and learning outcome (or cognitive performance) measures by determining whether an instructional manipulation that increases learning scores also increased eye-fixation times. The first two lines of Table 3 involve two research projects on signaling effects (Boucheix & Lowe, 2010; De Koning et al., 2010) in which visual signals improved learning outcome scores (or measures of deep cognitive processing during learning) and improved fixation time scores. Thus, in both cases there is evidence of a link between increases in the perceptual processing of relevant portions of the graphic and improvements in measures of learner understanding of the material.

The second set of two lines of Table 3 involve two research projects on prior knowledge effects (Canham & Hegarty, 2010; Jarodzka et al., 2010) in which people with higher prior knowledge performed better on eye-fixation time scores and cognitive performance scores than did people with lower prior knowledge. Thus, in both cases there is evidence of a link between increases in perceptual processing of relevant portions of the graphic and improvements in measures of cognitive performance on an intellectually demanding task.

The final two lines in Table 3 involve research projects on modality effects (Schmidt-Weigand et al., 2010) and presentation pacing effects (Meyer et al., 2010), respectively. However, neither project produced a strong effect on learning outcomes so it is not possible to determine whether there is a link between a manipulation having an effect on learning outcome and having an effect on eye fixations.

Overall, in studies where a manipulation had an effect on learning outcomes it also had a corresponding effect on eye-fixation time. Thus, this special issue contains several theoretically important instances of the link between looking and learning, as summarized in the first four rows of Table 3.

Table 3
Is there a link between learning outcome and perceptual processing during learning (or between cognitive performance and perceptual processing during performance)?

Study	Principle	Learning outcomes	Perceptual attention	Link
Boucheix and Lowe, 2010	Signaling	Yes	Yes	Yes
De Koning et al., 2010	Signaling	Yes	Yes	Yes
Jarodzka et al., 2010	Prior knowledge	Yes	Yes	Yes
Canham and Hegarty, 2010	Prior knowledge	Yes	Yes	Yes
Schmidt-Weigand et al., 2010	Modality	No	Yes	No
Meyer et al., 2010	Pacing	No	No	No

5. Practical contributions

As summarized in Table 4, the papers in this special issue address four principles of instructional design for online instructional graphics, that is, the signaling principle, the prior knowledge principle, the modality principle, and the pacing principle. The signaling principle calls for adding cues that highlight the relevant portions of graphics. Although prior research has sometimes failed to find support for the use of visual cues such as arrows in animations (Boucheix & Lowe, 2010; Kriz & Hegarty, 2007; Mautone & Mayer, 2001), the authors in this special issue have identified two versions of visual signaling that improve learning outcomes or cognitive processing during learning: (a) *spreading color cues* and (b) *shading of irrelevant features*. Boucheix and Lowe (2010) used spreading color cues that highlighted the relevant features of an animation depicting how a piano works; the colorization occurred in temporal correspondence to events in the animation and used different colors for different subsystems. In this way, the visual cues conveyed information spatially and temporally. In particular, spreading color cues provided a progressive pathway highlighting the causal chain among components in the piano mechanism. De Koning et al. (2010) used shading of irrelevant features of an animation depicting how the human cardiovascular system works; a different

Table 4
Four instructional design principles for graphics.

Principle	Definition
Signaling	People learn better (or perform better) from graphics when relevant features are highlighted rather than not highlighted.
Prior knowledge	People with high prior knowledge learn better (or perform better) from graphics than do people with low prior knowledge.
Modality	People learn better from graphics that are accompanied by spoken text rather than printed text.
Pacing	People learn better from an animation played at high-speed followed by slow-speed than vice versa.

subsystem was central to each of five phases of the animation so the shading shifted to subsystems that were not the focus for each phase of the animation. As in the Boucheix and Lowe study, the visual cues used by De Koning et al. conveyed information both spatially and temporally. Thus, an important practical breakthrough in this work is that visual cues can be effective in promoting learning when they change over time and highlight the entire subsystem to be viewed.

The prior knowledge principle calls for providing students with relevant prior knowledge before they view an instructional graphic. The principle worked both in a between subjects design involving how biology professors and students make judgments about locomotion of fish they see in an video (Jarodzka et al., 2010) and in a within subjects design involving how students make judgments about wind direction in weather maps before and after basic instruction in the causes of wind (Canham & Hegarty, 2010). This practical advice is similar to Mayer's pre-training principle (Mayer, 2005, 2009), which calls for providing pre-training to novices in the names and characteristics of key components in a causal system.

The modality principle (Ginns, 2005; Low & Sweller, 2005; Mayer, 2009) is the most well-established principle of multimedia design. It calls for supplementing an animation with spoken text rather printed text, particularly when the animation is presented at a fast pace under system control and the words are familiar to the learner. Although Schmidt-Weigand et al. (2010) did not find strong evidence for the modality principle (perhaps for reasons discussed above), the preponderance of evidence still supports the modality principle.

The pacing principle was not supported by Meyer et al. (2010, Experiment 2), nor does it have an extensive support in the literature. Thus, it is prudent to refrain from implementing the pacing principle until more conclusive supporting evidence is available.

6. Conclusion

The design of learning environments involving graphics should be consistent with a research-based theory of how people learn and evidence-based principles of how to help people learn. Research using eye tracking offers a unique path to testing aspects of theories of multimedia learning, particularly concerning perceptual processing during learning. The authors of this special issue are to be commended for their efforts to add to the evidence base on how people learn and think with graphics.

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruickshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Boucheix, J.-M., & Lowe, R. K. (2010). An eye-tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction, 20*(2), 123–135.

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*(2), 155–166.

Clark, R. C., & Mayer, R. E. (2008). *e-Learning and the science of instruction* (2nd ed.). San Francisco, CA: Pfeiffer.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from complex animation: seeing is understanding? *Learning and Instruction, 20*(2), 111–122.

DeLeeuw, K., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology, 100*, 223–234.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313–331.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: how experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146–154.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*, 911–930.

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147–158). New York: Cambridge University Press.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*, 377–389.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

Meyer, K., Rasch, T., & Schnotz, W. (2010). Effects of animation's speed of presentation on perceptual processing and learning. *Learning and Instruction, 20*(2), 136–145.

O'Neil, H. F. (Ed.). (2005). *What works in distance learning: guidelines*. Greenwich, CT: Information Age.

O'Neil, H. F. (Ed.). (2008). *What works in distance learning: sample lessons based on guidelines*. Greenwich, CT: Information Age.

Pelligrino, J. W., Chudowski, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: the science and design of educational assessment*. Washington, DC: National Academy Press.

Schmidt-Weigand, F., Kohert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20*(2), 100–110.