Learning and Instruction 20 (2010) 172–176

Learning and Instruction

www.elsevier.com/locate/learninstruc

Commentary

The use of eye movements in the study of multimedia learning

Jukka Hyönä*

Department of Psychology, University of Turku, FI-20014 Turku, Finland

Abstract

This commentary focuses on the use of the eye-tracking methodology to study cognitive processes during multimedia learning. First, some general remarks are made about how the method is applied to investigate visual information processing, followed by a reflection on the eye movement measures employed in the studies published in this special issue. It is argued that global eye movement measures indexing attentional and encoding processes during the entire learning period should preferably be complemented with more fine-grained analyses that are either time-locked to important events taking place in an animation or that by other means provide information about the time course of learning. As nicely documented in the present set of studies, it is also of importance to complement the eye-tracking data with offline measures indexing the end product of learning. Such a complementary approach is likely to yield important new insights into the process of multimedia learning.

Keywords: Eye movements; Multimedia learning; Visual attention; Comprehension

1. Introduction

In this commentary, my focus is on the use of the eye-tracking methodology to study cognitive processes during visual learning tasks, especially during learning from written texts, graphics and animations. I start out by saying that I very much welcome the application of the method to multimedia learning. As the papers of this special issue witness, eye-tracking can reveal important insights into the ongoing learning process. To date, the method has been successfully applied, for example, to the study of cognitive processes in scene perception (for a review, see Henderson, 2003) and in reading (for a review, see Rayner, 1998), but eye-tracking studies of the processing of multimedia materials are still relatively sparse.

In studies of scene perception, one key question has been to determine the relative contributions of low-level visual features versus higher-level cognitive factors on human gaze behavior. In reading research, the focus has been on uncovering the mental processes that contribute to successful recognition of words and to successful parsing of the sentence structure. However, relatively little effort has been devoted to the processing of expository texts mimicking textbook materials (for exceptions, see e.g., Hyönä & Lorch, 2004; Hyönä, Lorch, & Kaakinen, 2002; Hyönä & Nurminen, 2006; for possible eye movement measures to be used in such studies, see Hyönä, Lorch, & Rinck, 2003). Although I think it is a pity that eye-tracking has not been used much to study the processing of and learning from expository texts, I am also aware at least of some of the reasons for this shortage. One has to do with the fact that the method provides a very rich data set that may be highly challenging to analyze and make sense of. A similar challenge is faced by researchers applying the method to multimedia learning. However, as it becomes evident from this special issue, the challenge can be successfully met.

My commentary is structured as follows. First, I present some background to eye movement research for readers not familiar with the method, followed by a brief discussion of the limitations of the method. The primary focus, however, is on the choice of eye movement measures employed in the present set of studies. In this section I make suggestions for how such

* Corresponding author. Fax: +358 2 333 5060.
E-mail address: hyona@utu.fi

0959-4752/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.learninstruc.2009.02.013

data may be analyzed in further detail to tap into the time course of the learning process.

2. Some general notes on eye movement research

The present third era of eye movement research (see Rayner, 1978, 1998) is characterized by the use of human eye movements to index mental processes that are ongoing when people interact with different types of visual environments (e.g., written texts, illustrations, human facial expressions, or traffic scenes). In doing so, researchers have subscribed to the so-called *eye-mind hypothesis* (Just & Carpenter, 1980), according to which there is a close link between the direction of human gaze and the focus of attention. In other words, it is assumed that people attend to and process the visual information that is currently looked at. Naturally, in order to this assumption to hold, the available visual environment needs to be relevant to the task at hand. While I am writing this text, I find myself looking out of my office window at the library facilities located opposite to my office. However, my mind is busy finding a good formulation to my thoughts, so my gaze behavior is not reflecting what I am attending **to** at the moment. With this example I hope to illustrate the point that gaze behavior can serve as an index of current attentional processes only as long as the available visual environment in front of our eyes is pertinent to the task we would like to study.

The third era of eye movement research began soon after the cognitive revolution in psychology. The research was further boosted by the availability of microcomputers and commercially available recording apparatuses (see McConkie, 1997, for a reminiscence of the early days of eye movement research). Recent technological development has made the devices increasingly user-friendly: eye-trackers are relatively easy to operate, they are often unobtrusive to the participant, and ready-made analysis software packages greatly help to make sense of rich data sets. In the earlier days, researchers needed to write their own software to collect and analyze data; thus, the methodology was available only for the most devoted ones.

When people interact with visual environments, they make a sequence of fixations separated by fast eye movements (so-called saccades that are the fastest motoric movements human beings can make). Intake of visual information takes place during fixations, while saccades bring the center of the eyes (fovea) to new locations in the visual scene. Depending on the visual task and the momentary processing difficulty, individual fixations typically last about 200−500 ms. There exists now ample evidence demonstrating that increased processing difficulty is capable of inflating the duration of individual fixations. Moreover, fixation density may also be affected (i.e., the number of fixations is greater on a stimulus that is difficult to encode and/or comprehend). As evidenced by the present set of studies, also relevance assignment influences eye behavior. More fixation time is devoted to task-relevant stimulus features than task-irrelevant features (see, e.g., Kaakinen, Hyönä, & Keenan, 2002, for a similar finding in the processing of expository texts). Finally, as rightfully pointed out in several papers, the eyes are also drawn to visually salient features in our environment. For example, abrupt onset of stimulus, motion, and stimulus brightness are features capable of attracting the eyes. In sum, the eyes are guided both endogenously (i.e., to meet the learner's task-relevant goals) and exogenously (i.e., by perceptually salient stimulus characteristics). The present set of papers is to be praised for taking seriously into consideration both types of factors (see also, e.g., Lowe, 2003).

3. Limitations of the eye-tracking method

It is important to note the limitations of the eye-tracking method. Even though it provides highly valuable information (i.e., about what is perceived as task-relevant) it does not as such tell the researcher anything about the success or failure of comprehending the relevant piece of information. The learner may spend a lot of time attending to a relevant stimulus feature without adequately comprehending its relevance or the underlying principle it denotes (e.g., the learner may be looking at visually cued features in an animation of the workings of the human cardiovascular system without necessarily comprehending the operation of the cued subsystem; see De Koning, Tabbers, Rikers, & Paas, 2010). Thus, the eye-tracking data must be complemented with other performance measures, such as retrospective comprehension tests or think-aloud protocols. The studies included in the special issue are excellent examples of this complementary approach (see also Kaakinen & Hyönä, 2005, for an example of such an approach applied to the study of text comprehension).

4. Complementing eye-tracking with offline measures

An interesting and innovative approach in combining eye-tracking with a retrospective report on the comprehension process is used by (Jarodzka, Scheiter, Gerjets, and Van Gog, 2010; see also De Koning et al., 2010). They showed the learners their eye movement patterns registered when they viewed videos of different types of swimming fish in order to determine the fish's locomotion pattern. Using the eye movement pattern as a memory cue learners were asked to verbally report their thought contents while they were viewing the animation. The idea is that the played-back eye movement pattern may cue the learners to recover how they encoded and interpreted the various stimulus features. This sounds an intriguing approach definitely worth trying out in future studies. To further improve the effectiveness of the cue, perhaps one could use this "cued retrospective reporting" after each stimulus presentation, rather than at the end of the study. The downside of this could be that learners become increasingly aware of their eye movements, which may influence the viewing of subsequent animations. On the positive side, a frequent exposure to learners' own eye movement patterns may increase their intrinsic value as an effective memory cue.

Playing back eye movement data to the participants may also be an informative and useful tool to teach efficient