Chapter 33
The Use of Eye-Gaze to Understand Multimedia Learning

Malinda Desjarlais
Mount Royal University, Canada

ABSTRACT

There is great interest in how learners construct knowledge when presented with multimedia. Although information has been gained from evaluating recall performance across multimedia conditions, the use of eye-gaze indices for understanding multimedia learning is becoming increasingly popular. Within the multimedia learning literature, researchers have used duration, frequency and sequence of fixations as well as shifts in eye-gaze to identify differences regarding the selection, organization, and integration of information among learners. The current chapter provides a discussion of eye-gaze measures that have been employed in multimedia research and their related interpretations for the attentional process that takes place during the learning phase. In addition, considerations for using eye-gaze measures to understand multimedia learning are presented.

INTRODUCTION

Educators and researchers have long been interested in how learners construct knowledge. Relatively recently, attention has been devoted to information processing when learners are presented with multimedia. Multimedia presents learners with information consisting of verbal (spoken and/or written text) and pictorial (static and/or dynamic visualizations including drawings, photos, animations, graphics, and videos) representations which may or may not be relevant to the individual's learning goals. Multimedia presentations are increasingly being used to present information in both formal and informal learning contexts. These contexts, include but are not limited to the traditional classroom, online during distance learning courses, slideshows, Internet websites, educational-focused apps on smartphones and tablets, and during work presentations. Much empirical research supports that people learn better from multimedia, when the presentation is designed effectively, than from words alone (Fletcher & Tobias, 2005).

DOI: 10.4018/978-1-5225-3822-6.ch033

In effect, there has been a call for investigations of information processing during multimedia learning and the variables that influence knowledge acquisition.

Although information has been gained from evaluating recall performance across multimedia conditions (see Mayer, 2005 for a review), the use of eye-movement indices for understanding multimedia learning is becoming increasingly popular. Indeed, area journals have devoted special issues to eye-gaze research regarding multimedia learning (e.g., ACM Transactions on Multimedia Computing, Communication & Applications, 2007, vol. 3; Learning & Instruction, 2010, vol. 20). To highlight the unique knowledge eye-gaze recordings have contributed to the understanding of learning from multimedia, the current chapter is structured as follows: background to eye-gaze research is provided for readers not familiar with the methodology, followed by a discussion of information processing during multimedia learning based on eye-gaze measures. The objective of the current chapter is not to provide a comprehensive review of the multimedia literature, but rather to provide a discussion of the eye-gaze measures that have been employed in multimedia research and their related interpretations for cognitive processing during the learning phase.

BACKGROUND ON EYE-GAZE RESEARCH

According to the eye-mind hypothesis (Just & Carpenter, 1980) and the assumption of the eye-mind link (Reichle, Pollatsek, & Rayner, 2006), there is a close association between the element of the visual field a person is fixating on (aka overt attention) and his/her cognitive attention (aka covert attention). It is assumed that individuals are attending to the element of the visual field which their eyes are fixated (Anderson, 2000; Wang, 2011). When people attend to information they make a sequence of fixations separated by fast eye movements known as saccades. Fixations capture the motionless gaze or pauses that typically last about 200-500ms when visual information enters the information processing system; on the other hand, vision is suppressed during saccades (Rayner, 1998; Viviani, 1990). Therefore, information abstracted from learners' fixations - including duration, frequency, location, or sequence - has been considered the most meaningful for understanding attentional processes during multimedia learning.

According to Lai et al. (2013), popular eye-gaze measures can be categorized into the following three different scales of measurement: temporal, count, and spatial. Temporal eye-gaze measures involve the length of time a person is fixated on a particular element of the visual field, either on average, in entirety, or relative to other fixation locations; whereas, count measures involve frequency counts of fixations in particular areas of the visual field. While the most popular method for assessing differences in frequency and duration of fixations has been quantitative, some researchers have compared heat maps to assess temporal differences among learners. Heat maps show the overall visual distribution using color codes to differentiate the time spent fixating on each location of the visual field. This visual display enables researchers to identify at a quick glance which elements were attended to most and least often for either individuals or on average when collapsed across participants in a particular group/condition. Often the number and length of fixations have been used to answer questions related to how learners process information and which elements in the visual field receive greater attention.

Generally, it is agreed upon that fixation frequency and duration indicate attention (Wang, 2011) and the amount of cognitive processing engaged with fixated information (Anderson, Bothell, & Douglass, 2004; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Just & Carpenter, 1980; Rayner, 1998). However, it is ambiguous as to whether extensive processing of a particular element is due to perceived impor-