Animation and narration: Using eye tracking to understand visual attention distribution

Eric N. Wiebe & Leonard A. Annetta
Dept. of Mathematics, Science and Technology Education, North Carolina State University, Box 7801, Raleigh, NC, 27695-7801 USA
Email: eric_wiebe@ncsu.edu; len_annetta@ncsu.edu

Presented at AERA 2007
Chicago, IL, USA

Abstract

Research and application of cognitive load theory has demonstrated the effect of various text/graphic/narration relations on learning using multimedia material. Related work has looked at how the degree of integration between the text and graphics influences their use in learning. This study set out to look at how the degree of integration between text and graphics interacts with graphic dynamics and narration to influence visual attention in multimedia instructional material. Eye tracking methodologies were used to explore how visual attentional resources were distributed under these varying conditions. Narration had a clear effect of "pacing" the viewers of the slide show. Static and animated graphics that were not relevant to the printed text (low integration) favored more gaze time on the text, though narration did tend to shift attention back to the graphic. In addition, animation seemed to create high load conditions that favored students relying more heavily on narration, shifting visual attention away from the text and to the graphic, especially in the high integration condition.

Summary

An important area of multimedia research has been on the use of dynamic graphics (animations). Researchers have reported on the powerful combination of animations and narration as an instructional design strategy (Mayer, 2003; Rieber, Tzeng, & Tribble, 2004). The advantage of this combination is based, in part, on both cognitive load theory (Paas, Renkl, & Sweller, 2003; Sweller, Merrienboer, & Paas, 1998) and dual coding theory (Paivio, 1986) However, research has been conflicted as to the advantage of using animations for conceptual learning (Mary Hegarty, 2004; Hutcheson, Dillon, Herdman, & Wood, 1997).

While both static and dynamic (animated) graphics demand visual attentional resources, animations have increased cognitive demands of animations over static graphics (Bodemer, Ploetzner, Bruchmuller, & Hacker, 2005; Lowe, 1999, 2003). Narration has been investigated as an approach to mediating cognitive load (Mayer & Moreno, 1998; Tindall-Ford, Chandler, & Sweller, 1997) by leveraging the distinct processing mechanisms for visual and auditory information (Baddeley, 1999; Paivio, 1986).

When looking at the distribution of visual attention between text and graphic elements, a factor of interest is the semiotic relationship between the textual and graphic content (Carney & Levin, 2002). How well these two elements are mutually beneficial will be determined in part by whether the two sources of information link as a unified conceptual idea. Research by Levin and others has demonstrated that poor integration of

text and graphics can be detrimental to learning with either the text or the graphic being reduced to simply being a "distraction"—that is, capturing visual attention that could otherwise be directed at information sources useful to the task at hand.

While measurement of learning outcomes has been the primary method of investigating these instructional design factors, a potential method for investigation is through the use of eye tracking technologies (M. Hegarty, Carpenter, & Just, 1991; Pellegrino, Chudowsky, & Glaser, 2001). Past research have demonstrated the linkage between eye fixation behavior (locus, duration, and sequence) and a cognitive processing model for both graphic and textual visual materials (Just & Carpenter, 1976).

The line of questioning addressed in this research is how eye tracking methodologies can be used to explore how narration, text and animations interact to influence the distribution of visual attentional resources across instructional materials. These findings, it is hoped, will point to ways in which cognitive load can be optimized in instructional media and, in turn, lead to improved learning outcomes.

Methods and Materials

Undergraduate students (N=22) were asked to view a 20-slide PowerPoint presentation on work and simple machines. All of the slides contained both a block of text and a graphic. They were told that they would be tested on the content at the end of the slide show. Half the students viewed the slides with an accompanying audio narration (Narration group) while the other half did not (Silent group). The narration paralleled the content of the text but did not read it word for word. For both groups, viewing was self-paced.

An IR-based combined pupil and corneal reflection eye tracker was used to measure the eye location (gaze) and fixations occurring within the defined look zones of graphic and text. Slides were chosen further analysis that varied in degree of information integration between the text and graphic and whether the graphics were static or animated. Slide HS (31 words; narration 36 sec) and Slide LS (28 words; narration 18 sec) used static graphics, while Slide HA and Slide LA (17 and 15 words; 16 and 7 sec, respectively) used animations. As a second factor, the graphics in Slides 3 and 9 were highly integrated with the accompanying text, while Slides 5 and 12 were not.

Results

For Slide LS (Low Integration, Static Graphic), there was a significant difference in the total gaze time between the Silent (M=12.6 sec) and Narration groups (M=17.7 sec) (F=8.17, p<.01). While the text-to-graphic fixation ratio for the Silent group was higher (M=8.23) than the Narration group (M=4.78), the difference did not reach significance (F=2.26, p<.15).

The animation slides showed slightly different trends than the previous slides. For Slide HA (High Integration, Animation), there was a significant difference in the total gaze time spent on the slide between the Silent (M=11.24 sec) and Narration groups (M=16.45 sec), (F=12.15, p<.0025). Like Slide HS, the text-to-graphic fixation ratios for both groups were quite low, but the Silent group was significantly higher (M=2.30) than the Narration group (M=0.59), (F=6.42, p<.02).

Unlike for any of the other slides, Slide LA (Low Integration, Animation) there was no significant difference in the total gaze time spent on the slide between the Silent (M=14.7 sec) and Narration groups (M=17.2 sec). The text-to-graphic fixation ratio for

the Silent group (M=9.92) was higher than the Narration group (M=5.47), however this difference did not rise to the level of significance (F=1.40, p<.25).

Also of note, Slides LS and LA showed higher text-to-graphic ratios when there was no Narration, with Slide LA (Animation) showing a significant difference. Viewers of Slide HA (Animation), when Narration was present, spent more time on the graphic than the text.

Educational Importance

In general, the predictions of the effect of media manipulations made by cognitive load theory are supported by the eye tracking data. Narration can have a significant impact on how attentional resources are divided between the text and graphics regardless of whether the graphic is static or dynamic. In both this study and in Slykhuis, Wiebe and Annetta (2005), narration had a clear effect of "pacing" the viewers of the slide show. In addition, when paired with highly integrated text, relevant animation seemed to create high load conditions that favored students relying more heavily on audio narration. While the eye tracking data does not directly correlate to higher retention or comprehension of material, future work with this data set will look into just this question.

References

Baddeley, A. (1999). *Human memory*. Boston, MA: Allyn & Bacon.

Bodemer, D., Ploetzner, R., Bruchmuller, K., & Hacker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science, 33*, 73-95.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(1), 5-26.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14*(3), 343-351.

Hegarty, M., Carpenter, P. A., & Just, M. A. (1991). Diagrams in the comprehension of scientific text. In R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 641–668). New York: Longman.

Hutcheson, T. D., Dillon, R. F., Herdman, C. M., & Wood, J. (1997). *To animate or not to animate, that is the question.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 41st annual meeting - 1997, Albuquerque.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441-480.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225–244.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction, 13*(2), 157-176.

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction, 13*(2), 125-139.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1-4.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

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

Rieber, L. P., Tzeng, S.-C., & Tribble, K. (2004). Discovery learning, representation, and explanation within a computer-based simulation: Finding the right mix. *Learning and Instruction, 14*(3), 307-323.

Slykhuis, D. A., Wiebe, E. N., & Annetta, L. A. (2005). *Eye tracking students' use of science related Powerpoint presentations.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Dallas, TX.

Sweller, J., Merrienboer, J. J. G. v., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology-Applied, 3*(4), 257-287.