Students regulate their learning processes as a function of multimedia coherence: Analyses of eye-gaze behaviour

Gregory Trevors, Reza Feyzi-Behnagh, Anoop Saxena, François Bouchet, and Roger Azevedo

McGill University, Canada

{gregory.trevors@mail.mcgill.ca; reza.feyzibehnagh@mail.mcgill.ca; anoop.saxena@mail.mcgill.ca; francois.bouchet@mcgill.ca; roger.azevedo@mcgill.ca}

Abstract

This study investigated students' regulation of learning processes across science multimedia. Within an adaptive, hypermedia learning environment, 81 university students were presented 38 pages of texts and images of the circulatory system. A subset of 4 pages were rated for high or low coherence between text and image. Participants' learning strategy behaviours (i.e., seeking out, attending to, and integrating relevant multimedia content) were assessed by aligning eye-tracking and computer-based learning environment log-files. Subsequent learning achievement was assessed with a multiple choice posttest. Result show that students regulated their learning processes as a function of text-image coherence, which was related to learning achievement: processes on high-coherence pages positively related to learning, whereas processes on low-coherence pages negatively related to learning. Implications for theory and educational practice are discussed.

Aims

Effective computer-based learning environments are those that promote the learner's use of cognitive processes integral to quality learning and reduce other processes that distract from these efforts. Beneficial processes include seeking out, attending to, and integrating relevant multimedia content with the ultimate goal of constructing a coherent mental representation of instructional content. Students who regulate their learning to deploy these strategies for relevant multimedia content achieve better learning outcomes than their counterpart who do not engage in self-regulated learning (Azevedo et al., 2012). However, these efforts are hindered with the inclusion of content superfluous to comprehension, as processing extraneous content competes for students' limited cognitive resources (i.e., the coherence principle; Mayer, 2009). It is therefore pertinent for educational researchers to obtain behavioural data on students' use of adaptive learning strategies. Thus, the purpose of this study was to investigate how students regulated their cognitive processes — that is, seeking out, attending to, and integrating relevant multimedia content — as evinced through eye-gaze analysis, and to determine which processes relate to high quality learning. We made two hypotheses: first, eye-gaze patterns would differ as a function of the degree of coherence between text and images; second, processing associated with low-coherence content would negatively relate to learning achievement.

Method

Participants

A total of 81 undergraduate students, recruited from a large university in North America, took part in this study (N = 65 females), with a mean age of 21.3 years (range of 18-31 years), mean GPA of 3.27 (SD = .42).

MetaTutor

MetaTutor is a multi-agent, hypermedia, computer-based learning environment that contains 38 pages of text and images on the human circulatory system (see Azevedo et al., 2012).

Procedure

Participants had one hour to learn as much as they could about the circulatory system. When they navigated to a page, only text was initially visible, and one image appeared as a miniature thumbnail with the annotation "Click to enlarge image." Participants had the choice whether to click on the image to view it fully.

Data Sources

To address our hypotheses, product (i.e., 25 multiple choice posttest questions) and process data (i.e., eye-tracking and log-files of human-computer interactions) were collected from participants as they learned with MetaTutor.

Independent variable. Coherence between text and images was determined for pages of MetaTutor content. Pages were reviewed and sorted into three broad categories for the degree of semantic overlap between texts and images (i.e., high, moderate, and low semantic relation). Two pages from each of the low- and high-coherence categories were selected to achieve maximum variation for analysis.

Dependent variables. Timestamps for when participants navigated to these pages and opened an image were extracted from MetaTutor log-files, which were used to extract relevant data from eye-tracking log-files. Areas of interest (AOIs) were overlaid on text and image boxes in MetaTutor's interface to collect gaze behaviour data at those areas. From these data sources, five dependent variables were extracted: first, the proportion of images that were processed (PIP) was operationalized as the average proportion of images on high- and low-coherence pages that were opened by participants. Second and third, the durations in seconds of fixations on text and image AOIs were recorded. Fourth, participants' use of the strategy coordinating of informational sources (COIS) was operationalized as a sequence of two fixations between text and image, the frequencies of which were tallied. Last, proportions of correctly answered posttest questions were extracted, for all questions and questions specific to pages visited.

Findings

Four ANOVA tests were conducted on four of the dependent variables with high- or low-coherence entered as the independent variable. Not all participants navigated to all four pages, thus ruling out a single MANOVA analysis. Descriptive and inferential statistics are contained in Tables 1 and 2. Overall, participants exerted greater cognitive effort on high-coherence pages, evinced by choosing to view a greater proportion of such images, longer reading and image inspection times, and greater frequency of integrative behaviours, compared to low-coherence pages.

Additionally, correlations between eye-gaze and learning outcomes were obtained. For scores on all questions in the posttest, COIS on high-coherence pages showed a positive correlation, $r(35) = .34$, $p > .05$. For scores on posttest questions related to only those pages participants visited during their session, PIP on low-coherence pages showed a negative

correlation, $r(69) = -.30$, $p < .01$, whereas reading time across pages showed a positive correlation, $r(42) = .31$, $p < .05$.

Theoretical and Educational Significance

The current study showed the theoretical and educational significance of coherence between text and images by analyzing patterns of eye-gaze in an hypermedia learning environment. Students in this study sought out a higher proportion of images for high- rather than low-coherence pages, indicating that they likely recognized that high-coherence pages contained difficult text and chose to supplement their comprehension with visual information. The decision to seek out more images on high-coherence pages was linked to increased processing and integration of text and image content, which were in turn related to subsequent learning achievement. Coupled with the finding that processing related to low-coherence pages was negatively related to learning, we can conclude behavioural support for the coherence principle (Mayer, 2009). Practically, these data can be used in the design of multimedia learning environments, specifically, by prompting students to integrate textual and visual information on high-coherence pages, and by eliminating occasions for students to expend precious cognitive processing on weakly coherent images.

References

Azevedo, R., Landis, R., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J., . . . Hossain, G. (2012). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. In S. A. Cerri et al. (Eds.), *LNCS: Vol. 7315. 11th International Conference on Intelligent Tutoring Systems* (pp. 212-221). Berlin, Germany: Springer-Verlag. doi: 10.1007/978-3-642-30950-2_27

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

Table 1

Descriptive statistics for eye-gaze patterns across high and low multimedia coherence groups

Dependent Variable	Low coherence			High coherence		
	M	*SD*	*N*	*M*	*SD*	*N*
PIP	.24	.39	70	.52	.38	70
Text fixation [a]	15.29	20.81	18	58.47	75.51	45
Image fixation [a]	6.14	5.05	20	32.59	54.28	47
COIS [b]	2.57	4.83	21	17.49	26.73	51

[a] time in seconds; [b] frequency count

Note: PIP = Proportion of images processed; COIS = Coordinating informational sources.

Table 2

Inferential statistics for eye-gaze patterns across high and low multimedia coherence groups

Dependent variables	ANOVA statistics			
	Error *df*	*F*	*p*	η_p^2
PIP	138	7.72	< .01	.05
Text fixation [a]	61	5.66	< .05	.09
Image fixation [a]	65	4.69	< .05	.07
COIS [b]	70	6.41	< .05	.08

[a] time in seconds; [b] frequency count

Note: PIP = Proportion of images processed; COIS = Coordinating informational sources.