Eye Tracking Students' Gazes on Feedback in a Digital Assessment Game

Maria Cutumisu, Krystle-Lee Turgeon, Lydia Marion González, Tasbire Saiyera, and Steven Chuong
cutumisu@ualberta.ca, krystlel@ualberta.ca, gonzleze@ualberta.ca, saiyera@ualberta.ca, schuong@ualberta.ca
University of Alberta

Daniel L. Schwartz, Stanford University, danls@stanford.edu

Abstract: This study tracked the eye movements of n = 24 undergraduate students while they played an assessment game, Posterlet. Students designed digital posters and then they received three pieces of constructive critical (negative) or confirmatory (positive) feedback on each of the posters. Total eye gaze duration analyses revealed that students spent significantly more time attending to the critical rather than to the confirmatory feedback they received. They also dwelled more on each word and letter of critical rather than confirmatory feedback. Finally, they also revisited critical feedback more often than confirmatory feedback. Implications of these results and future research directions are discussed.

Introduction

Feedback is one of the most impactful factors for learning (Hattie & Timperley, 2007). The feedback literature differentiates between positive (confirmatory) and negative (critical) feedback. Recently, eye movement research attempted to gain an insight into cognitive processes by tracking participants' eye gazes in real time and measuring two types of eye movements: fixations and saccades. Specifically, fixations are short gaze stops of approximately 200 milliseconds (ms) that are used to infer mindful cognitive processing (Bolzer, Strijbos, & Fischer, 2015), while saccades are fast gaze moving actions (Rayner, 1998). However, the eye tracking literature on feedback processes is scarce (Timms, DeVelle, & Lay, 2016). This research aims to gain an insight into the mechanisms of feedback processing by examining students' eye gazes on feedback while they play a digital game in which they design posters. Informed by prior research, this study hypothesizes that students dwell longer on critical than on confirmatory feedback when they are assigned feedback following a task.

Methods

A total of 24 university students (9 males and 15 females), ranging from 18 to 29 years of age, with a mean age of 21.79 years (SD = 3.1), took part in this study. Participants were recruited from a large North American university from a subject pool program. They received course credit for their participation. They signed an electronic consent form prior to joining the study and were tested individually in one session lasting approximately 45 minutes. The first 5 minutes were used to calibrate the eye tracker using a five-point calibration sequence. To calibrate and validate the eye tracker, participants had to follow a dot that appeared at five different locations on the computer screen. This procedure was repeated until the average deviation of the visual angle between the calibration was one degree. Participants played the Posterlet computer game for about 15-20 minutes. After completing the game, participants completed a post-test for about 20 minutes.

The study employed three instruments: (1) a computer-based assessment game, Posterlet (Cutumisu et al., 2015); (2) an eye tracker to capture students' gazes superimposed on the game; and (3) a post-test survey of background information, including demographic information. An alternative game version was designed as part of a larger yoked-study design. Players do not have a choice regarding the valence of their feedback in this version of Posterlet. Instead, they are assigned the feedback valence choices of participants who played the original Posterlet version. Eye movements were recorded using the SR Research EyeLink 1000 Plus desktop remote-mode system. The Screen Recorder software was employed to record participants' gazes onto the Posterlet game. Then, participants filled an online post-test survey.

Critical Feedback measures the number of critical feedback ("I don't like") messages that the participants encountered across the game. Gazes on Critical Feedback counts the number of critical feedback boxes where a participant's gaze was recorded across the three posters. Each of the critical feedback boxes was coded with 1 if there was a gaze ever detected on that box and with 0 otherwise. Mean Gaze Duration per Letter of Critical Feedback approximates the average time that participants spend looking at each letter of critical feedback across the Posterlet game. This measure is important, as it enables a fair comparison of the time participants took to attend to each feedback valence. Several steps are taken to compute this measure for each feedback valence. First, the sum of all the individual fixation durations on each feedback box, including the durations of the regressions on that box, is computed. Then, this measure is divided by the length (i.e., the number of letters, including spaces) of the feedback message in that box. Then, these values are added for all the boxes of each valence and divided by the Gazes on Critical Feedback to obtain an estimate of the average time spent per letter of feedback valence. Mean Gaze Duration per Word of Critical Feedback measures the average time a participant spent looking at each word of critical feedback across the game. Mean Number of Fixations on Critical Feedback represents the average number of a participant's gaze fixations on the critical feedback boxes across the game, ranging from 3 to 14. Mean Number of Regressions on Critical Feedback represents the average number of times a participant revisited the critical feedback boxes across the game, ranging from 0 to 2. These measures were further refined according to the length of feedback and the number of words of feedback.

Results

Do students spend more time actively looking at critical rather than at confirmatory feedback when feedback is assigned? The mean gaze duration across the game was significantly larger [t(20) = 4.93, p < .001] for critical (M=2387.42 ms, SD = 581.44 ms) than for confirmatory (M=1873.07 ms, SD = 637.74 ms) feedback. Analyses revealed that participants spent more time attending to critical feedback per letter [t(20) = 3.87, p < .01] and per word [t(20) = 3.67, p < .01] than to confirmatory feedback. On average, participants read critical feedback (M=.0062, SD=.003) at a slower pace [i.e., less words per millisecond; t(20)=-2.11, p=.048] than they read confirmatory feedback (M=.0085, SD=.005). The more the participants encounter critical feedback, the more they dwell on it per letter and per word, but there is no association between gazes on critical feedback and dwell time per letter and per word on confirmatory feedback.

Is there a difference in the mean number of gaze fixations on feedback between valences when feedback is assigned? A paired-samples t-test analysis showed that the mean number of fixations on critical feedback boxes (M = 9.37, SD = 2.47) was larger [t(20) = 5.97, p < .001] than the mean number of fixations on confirmatory feedback boxes (M = 6.71, SD = 1.95). This suggests an overall closer attention paid to critical than to confirmatory feedback. Participants also read critical feedback (M = .18, SD = .05) more closely per letter [t(20) = 4.87, p < .001] than confirmatory feedback (M = .13, SD = .04). They also read critical feedback (M=.85, SD=.23) more closely per word [t(20)=4.57, p < .001] than confirmatory feedback (M = .65, SD=.19).

Is there a difference in the mean number of feedback revisits between feedback valences when feedback is assigned? A paired-samples t-test analysis revealed that the mean number of regressions on critical feedback boxes (M=.85, SD=.61) was significantly larger [t(20)=4.10, p<.01] than the mean number of regressions on confirmatory feedback boxes (M=.40, SD=.41). This finding confirms the results of the previous analyses in this section, suggesting that, overall, participants attended to critical feedback boxes more often than to confirmatory feedback boxes. Participants also revisited critical feedback (M=.0157, SD=.01) more often per letter [t(20) = 3.93, p<.01] than confirmatory feedback (M=.01, SD =.01). They also revisited critical feedback (M=.08, SD=.05) more often per word [t(20)=3.89, p<.01] than confirmatory feedback (M=.04, SD=.04).

Conclusions and educational implications

Results suggest that students attended to critical feedback significantly more often and more closely than to confirmatory feedback. The dwell time (per letter and per word) was significantly larger for critical rather than for confirmatory feedback. This research may inform the design and the delivery of feedback, so that students could attend more to the type of feedback that helps them improve their outcomes the most.

References

Bolzer, M, Strijbos, J.W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: role of feedback-characteristics, fixation-durations and transitions. *Journal of Computer Assisted Learning, 31,* 422-434. doi:10.1111/jcal.12091

Cutumisu, M., Blair, K. P., Chin, D. B., & Schwartz, D. L. (2015). Posterlet: A game-based assessment of children's choices to seek feedback and to revise. Journal of Learning Analytics, 2(1), pp. 49-71.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Timms, M., DeVelle, S., & Lay, D. (2016). Towards a model of how learners process feedback: A deeper look at learning. *Australian Journal of Education, 60*(2):128-145.

Acknowledgments

We would like to thank the students who participated in this study, the University of Alberta *Support for the Advancement of Scholarship* Grant, and the *SSHRC IDG* Grant # RES0034954.