Human Factors in Training, Education, and Learning Sciences, Vol. 231, 2026, 115–120
https://doi.org/10.54941/ahfe1008002

Eye-Tracking Analysis of Students' Problem-Solving Behaviors for Learning Support in a Tutoring Context

Kento Yamao and Noriaki Kuwahara

Kyoto Institute of Technology, Kyoto, 606-0951, Japan

ABSTRACT

Understanding students' visual engagement during problem solving can provide insight into learning strategies beyond performance outcomes. This study explores gaze-based indicators of learning behavior in a real-world tutoring context using non-intrusive, webcam-based eye tracking. Junior high school students solved computer-based multiple-choice questions while gaze data were synchronized with problem-solving logs. We propose a simple analytical framework based on areas of interest representing question statements, answer choices, and non-task regions, and compute descriptive gaze metrics capturing attention allocation, comparison behavior, and exploration patterns. Group-level analysis based on task accuracy suggests that higher-performing students tend to exhibit more frequent transitions between questions and answer choices and more diverse gaze movements, while time-based measures show limited differentiation. These findings indicate that qualitative gaze patterns can provide complementary insight for supporting tutoring instruction and reflective learning support in everyday educational settings.

Keywords: Eye tracking, Eye gaze, Eye tracker, Education

INTRODUCTION

Understanding learners' cognitive processes during problem solving has long been an important topic in educational research, as observable outcomes such as accuracy or response time alone provide limited insight into how learners allocate attention and compare information. Eye-tracking technology has therefore been widely employed to investigate such processes by capturing learners' visual behavior during task execution.

Asakawa, Okano, and Hayashi analyzed gaze behavior during question answering using a dedicated eye tracker (Tobii), examining differences in responses to on-screen questionnaires as well as relationships among gaze patterns, questionnaire responses, and answer accuracy. Their study demonstrated the potential of eye-movement analysis for understanding learning-related behavior, while also relying on specialized hardware and controlled experimental settings (Asakawa, Okano, and Hayashi, 2020).

More broadly, advances in information technology have enabled the recording and analysis of non-instrumental human behaviors across multiple modalities, including gaze, gestures, and facial expressions. Such multimodal

Received February 5, 2026; Revised March 30, 2026; Accepted April 14, 2026; Available online July 20, 2026

behavior analysis has expanded beyond communication studies and has increasingly been applied to education, giving rise to learning analytics that record and analyze learners' activities and utilize the results in various forms of educational support (Owatari, Shimada, Minematsu, and Taniguchi, 2020).

In recent years, the COVID-19 pandemic has accelerated the adoption of online and hybrid learning environments worldwide. Within this context, methods for recording and analyzing learners' activities in online settings have gained increased attention, particularly as part of multimodal learning analytics frameworks (Katsumata, and Kato, 2019). These developments highlight the need for learning analysis techniques that can be applied in everyday learning environments without imposing additional burdens on learners.

Related research has also explored the extraction of physiological information from RGB camera images. For example, Poh et al. and Tsumura proposed systems for extracting pulse wave information from video recordings and estimating emotional states based on remote photoplethysmography. While these studies demonstrate the feasibility of non-contact sensing using generic cameras, they are not primarily designed for learning analytics, and further investigation is required to assess their applicability in real educational contexts (Poh, McDuff, and Picard, 2010, and Tsumura, 2021).

Building on these lines of research, the present study explores the use of webcam-based eye tracking as a practical and scalable approach for analyzing learners' gaze behavior in a real-world tutoring environment. Rather than focusing on fine-grained physiological estimation or laboratory-level accuracy, this work aims to examine whether coarse, interpretable gaze-based indicators can provide meaningful insight into learners' problem-solving strategies and support instructional reflection in everyday educational practice.

RELATED WORK

Eye tracking has long been employed to study visual attention and cognitive processes in reading, problem solving, and learning. In educational research, gaze metrics such as fixation duration, scan paths, and transition patterns have been associated with comprehension, expertise, and strategy use. Studies have reported that high-performing learners tend to exhibit more systematic comparison behaviors and broader visual exploration when solving complex tasks.

Recent work in learning analytics has emphasized the importance of interpretable indicators that can support instructors' understanding of learners' behavior. While machine learning approaches have shown promise in performance prediction, their black-box nature often limits practical adoption in educational practice.

However, many prior studies rely on high-precision eye trackers and laboratory-based tasks, which limit scalability and ecological validity. Fewer studies have explored the potential of generic, webcam-based eye tracking systems in everyday tutoring environments. This study addresses this gap by focusing on descriptive, task-agnostic gaze metrics that can be computed from low-cost eye-tracking data and applied in real instructional contexts.

STUDY DESIGN AND DATA COLLECTION

1. Participants and Learning Context

Participants were junior high school students enrolled in a private tutoring school. The study was conducted during regular learning sessions, aiming to minimize disruption to normal instructional activities. All data were collected with consent, and participation did not affect students' academic evaluation.

2. Learning Tasks

Students solved computer-based multiple-choice questions presented on a standard display. The tasks were designed to reflect typical tutoring materials rather than experimental stimuli. Although the present analysis focuses on multiple-choice questions, the study was designed with future extension to written and open-ended problems in mind.

3. Gaze Data Collection

Gaze data were recorded using a webcam-based eye-tracking system. This approach enabled non-intrusive data collection without requiring specialized hardware or calibration-intensive procedures. Gaze coordinates were recorded continuously and synchronized with problem-solving logs, including question timing and response events.

GAZE METRICS AND ANALYTICAL FRAMEWORK

1. Area of Interest (AOI) Definition

To analyze visual attention in a task-agnostic manner, the screen was divided into predefined Areas of Interest (AOIs):

- **Question area (Q):** containing the problem statement
- **Choice area (C):** containing the answer options
- **Non-task area:** regions unrelated to the task

This AOI structure was chosen to remain consistent across tasks and support future extensions beyond multiple-choice formats.

2. Gaze Metrics

Based on the AOI framework, the following metrics were computed:

- **ratio_Q:** proportion of total fixation duration on the question area
- **ratio_C:** proportion of fixation duration on the choice area
- **bounce_density_qc:** frequency of gaze transitions between question and choice areas
- **entropy_aoi:** entropy of AOI transitions, representing the diversity of gaze movement
- **response_time:** median response time per question

These metrics are intended as descriptive indicators of visual strategies rather than direct predictors of performance. In particular, they aim to capture how students distribute attention, compare information, and explore available options.

ANALYSIS PROCEDURE

Students were grouped based on task accuracy using a threshold-based criterion, resulting in High- and Low-performing groups. Group-level comparisons were conducted for each gaze metric to identify general tendencies rather than individual-level classification.

Statistical analyses were complemented by visualizations, such as boxplots and scatter plots, to emphasize interpretability and variance. This visualization-oriented approach aligns with the study's goal of supporting instructional reflection rather than automated decision-making.

RESULTS

Across multiple gaze metrics, differences in visual behavior tendencies were observed between performance groups.

High-performing students tended to show higher values across several gaze metrics, including **ratio_C, bounce_density_qc,** and **entropy_aoi.** These patterns suggest more active comparison among answer choices and broader exploration of task-relevant regions. In contrast, **ratio_Q** showed relatively small differences between groups, indicating that time spent viewing the question statement alone does not strongly distinguish performance levels.

Response time also did not show a clear advantage for high-performing students. While individual variance was large, the results suggest that higher performance is not simply associated with longer or shorter response times.

Additionally, scatter analysis between accuracy and response time highlighted the limitations of time-based indicators, reinforcing the importance of qualitative gaze patterns in understanding learning behavior.

DISCUSSION

1. Interpretation of Gaze Patterns

The observed tendencies suggest that successful problem solving may be associated with how students compare and revisit information rather than how long they read the question. Frequent transitions between questions and choices may reflect hypothesis testing or verification strategies, while higher entropy may indicate flexible exploration rather than rigid viewing patterns.

It is important to note that these interpretations are descriptive rather than causal. Gaze behavior may reflect underlying strategies, task familiarity, or individual differences, and should be interpreted in context.

2. Educational Implications

From an applied perspective, gaze-based indicators offer opportunities for supporting tutoring instruction. Visualizing where students hesitate, over-focus, or fail to compare options may help instructors adjust explanations or identify misconceptions more efficiently. Furthermore, such visual feedback can facilitate communication with parents by providing intuitive representations of learning processes beyond test scores.

3. Methodological Considerations

Webcam-based eye tracking inevitably involves noise and lower spatial precision compared to dedicated devices. However, the use of coarse AOIs and aggregated metrics mitigates some of these limitations. Rather than precise fixation analysis, the focus on relative attention distribution and transition patterns supports robustness and scalability.

FUTURE WORK

Future research will extend this framework in several directions. First, applying the same metrics to written and open-ended problem formats will help evaluate task generality. Second, longitudinal analysis across multiple sessions may reveal how gaze strategies evolve with learning. Third, integration with instructional design and real-time feedback systems could enable adaptive tutoring support.

Ultimately, gaze-based analytics may contribute to a broader ecosystem of learning support tools that emphasize transparency, interpretability, and collaboration among students, instructors, and parents.

CONCLUSION

This study presented a gaze-based analytical framework for examining learning strategies in a real-world tutoring context using webcam-based eye tracking. The results suggest that qualitative patterns of visual exploration provide complementary insights beyond traditional time-based measures. By focusing on interpretable metrics and practical applicability, this work highlights the potential of gaze-based learning analytics as a supportive tool for educational practice.

ACKNOWLEDGEMENT

The authors would like to thank the tutoring school for providing the learning environment and for their cooperation in this study.

REFERENCES

Asakawa, M., Okano, M., and Hayashi, H. (2020). *Analysis of response behavior to self-administered questionnaires using eye tracking*. Behaviormetrika, Vol. 47, No. 2, pp. 141–152.

Katsumata, H. and Kato, T. (2019). *Analysis of gaze data during online learning using eyeglass-type devices*. Proceedings of the 81st National Convention of IPSJ, Vol. 2019, No. 1, pp. 17–18.

Owatari, T., Shimada, T., Minematsu, T., and Taniguchi, R. (2020). *Development of a dashboard for real-time learning improvement based on learning logs of online electronic teaching materials*. IPSJ SIG Technical Report: Computer and Education (CLE), Vol. 2020-CLE-30, No. 14, pp. 1–4.

Poh, M., McDuff, D., and Picard, R. (2010). *Non-contact, automated cardiac pulse measurements using video imaging and blind source separation*. Optics Express, Vol. 18, No. 10, pp. 10762–10774.

Tsumura, N. (2021). *Affective engineering based on non-contact heart rate measurement from RGB camera video and its applications*. Oleoscience, Vol. 21, No. 5, pp. 159–166.