12th International Conference of the Immersive Learning Research Network
iLEAD Proceedings iLRN2026

DOI: https://doi.org/10.56198/j29ww513

Analysis of Attention Based on Gaze and Head Movements under Task Difficulty in an Extended Reality Diagnostic Simulation

Kukhyeon Kim[1[0000-0002-1542-176X]], Daeun Han[1[0009-0001-9208-8331]], and Jeeheon Ryu[1[0000-0003-2891-283X]]

[1] Chonnam National University, Gwangju, South Korea
jeeheon@jnu.ac.kr

Abstract. This iLEAD type's study examines how task difficulty shapes learners' visual and embodied exploration during strabismus diagnosis in an extended reality (XR) simulation. 21 medical students completed three diagnostic subtasks (ocular motility test, cover–uncover test, and prism cover test) while eye gaze and head movement data were continuously recorded (gaze hit position; head position and rotation). Visual exploration was quantified using gaze path rate (scanpath length per second) and gaze speed variability (SD of gaze speed), and embodied exploration was quantified using head translation path rate (head movement rate) and head angular speed variability (SD of head angular velocity). Behavioral indicators were compared across the three tasks (ocular motility test, cover–uncover test, prism cover test), accounting for repeated observations within individuals. Results showed that higher task difficulty was associated with increased gaze path rate, increased gaze speed variability, and increased head movement rate. Head rotation variability showed an increasing tendency with task difficulty but did not reach statistical significance. By integrating continuous eye and head movement indicators, this research provides quantitative evidence that task difficulty systematically alters how learners explore and physically navigate XR diagnostic simulations, offering a behavioral basis for understanding performance processes in immersive clinical training environments.

Keywords: XR Simulation, Gaze Movement, Head Movement, Attention, Task Difficulty.

1 Introduction

XR-based diagnostic simulations have gained attention as an alternative approach to clinical education because they enable safe, repeatable practice of procedures and diagnostic cues that resemble real clinical settings [1–4]. In these environments, however, learners must process complex visual information while navigating a three-dimensional space, manipulating tools, and making diagnostic decisions. Because these demands unfold in real time, diagnostic performance in XR is shaped not only by whether the learner succeeds, but also by how the learner explores the environment and coordinates visual and bodily actions during the task.

A key challenge in XR diagnostic training research is that many studies still rely heavily on post-task outcomes (e.g., accuracy or achievement scores). Outcome-focused evaluation provides limited insight into how learners allocate attention and how exploratory behaviors emerge moment by moment while the diagnosis is being performed [5]. To better understand the process of diagnostic reasoning in XR, it is necessary to use objective behavioral indicators derived from continuous tracking logs.

Eye-tracking metrics provide a direct window into how learners visually search for task-relevant cues. For instance, scan path-related measures can indicate the amount and extent of visual exploration, whereas variability in gaze speed may reflect unstable exploration dynamics or fluctuations in information seeking [2, 6]. Importantly, visual exploration in XR is tightly coupled with head and body movements because the user's viewpoint is physically situated in space. Therefore, movement-based indicators such as head translation (spatial exploration) and head rotation variability (instability in orientation dynamics) should be considered alongside gaze metrics to capture embodied exploration during XR diagnostic tasks [7–9].

From this perspective, differences in task difficulty systematically shape how learners explore and attend in XR diagnostic simulations. In strabismus diagnosis training, the required perceptual cues and procedural demands differ across tasks (ocular motility test, cover–uncover test, prism cover test), which can lead to distinct patterns of visual search and physical exploration. Accordingly, this study examines whether gaze- and head-based behavioral indicators vary as task difficulty increases across the three diagnostic tasks. We analyze gaze path rate and gaze speed variability as indices of visual exploration, and head translation path rate and head angular speed variability as indices of embodied exploration. By integrating these indicators, this research aims to provide a quantitative account of how learners' exploratory behaviors change across tasks in XR-based strabismus diagnosis simulations. In XR-based clinical training, there is growing interest in leveraging multimodal behavioral logs (e.g., gaze and head movement) to explain learners' performance processes beyond post-task outcomes [10].

- RQ1: Do gaze-based visual exploration indicators (gaze path rate, gaze speed variability) differ across tasks with increasing diagnostic demands (ocular motility test < cover–uncover test < prism cover test) in an XR simulation?
- RQ2: Do head-movement–based embodied exploration indicators (head translation path rate, head angular speed variability) differ across tasks with increasing diagnostic demands (ocular motility test < cover–uncover test < prism cover test)?

2 Method

2.1 Participants

Participants were 21 medical students (13 males, 8 females) enrolled in an ophthalmology clinical practice course. All were third- or fourth-year students at a national university medical school located in a metropolitan city and had completed relevant coursework and practical training related to strabismus. Recruitment was conducted via on-campus advertisements, and participants received a small participation incentive. The study was approved by the Institutional Review Board[1].

2.2 Experiment

Before the experiment, participants completed approximately 10 minutes of practice to familiarize themselves with the HoloLens 2 interface and object manipulation. Eye-tracking calibration was then performed to ensure accurate gaze data collection. The experimental setup is illustrated in Fig 1. Each participant completed three clinical subtasks for strabismus diagnosis: an ocular motility test, a cover–uncover test, and a prism cover test. The progressive difficulty of these tasks is a well-established standard in the regular strabismus training curriculum and was further validated by clinical specialists. In the ocular motility test, participants inspected extraocular movement patterns while following target-guided eye positions to judge abnormal motility; a representative in-simulation scene is shown in Fig 2. In the cover–uncover test, participants evaluated ocular alignment by covering one eye and observing refixation movements to identify manifest deviation. In the prism cover test, participants quantified the deviation by selecting and applying an appropriate prism while performing a cover procedure, which requires iterative adjustment and confirmation.

Fig. 1. Experiment Scene.

[1] IRB Approval No: 1040198-221021-HR-128-02

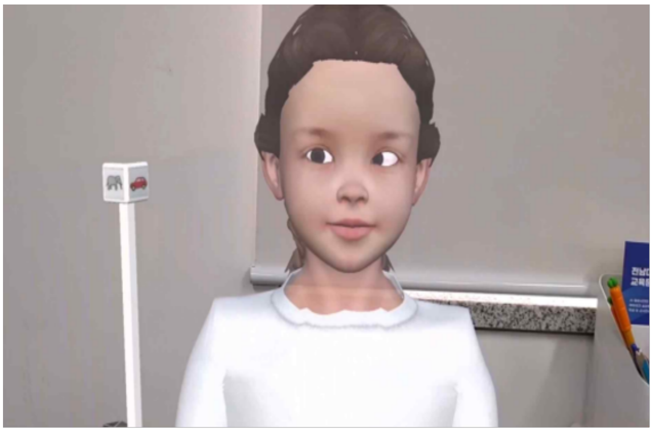

Fig. 2. Ocular Motility Test Scene.

During each trial, behavioral data were recorded continuously, including gaze tracking variables and head position and rotation. To objectively reflect the varying procedural demands, trials were strictly limited to 3 minutes for the ocular motility and cover-uncover tests, and 5 minutes for the prism cover test. Data collection ended when the participant indicated that the diagnosis was complete or when this time limit was reached. After each trial, participants reported their diagnostic judgment, including the presence and type of strabismus.

2.3 Variables

Table 1 summarizes the time-stamped tracking variables collected from the XR device. For head movement, head position (head_position_x/y/z) indicates the 3D location of the user's head center at each sample (left–right, up–down, forward–backward). Sample-to-sample changes in head position were accumulated to compute head translation path length, reflecting overall spatial exploration. Head rotation (head_rotation_x/y/z/w) captures head orientation in quaternion form. By computing the relative rotation between consecutive quaternions, we derived the rotation angle and angular speed; the standard deviation of angular speed was used as an index of head rotation variability.

For gaze, eye tracking was recorded using a ray-casting approach. Gaze origin (gaze_origin_x/y/z) is the starting point of the gaze ray, and gaze direction (gaze_direction_x/y/z) is its 3D direction vector. Gaze hit position (gaze_hit_position_x/y/z) is the intersection point between the gaze ray and the virtual scene, representing the 3D location of where the user was looking. Sample-to-sample distances between hit positions were accumulated to estimate gaze travel distance (scan path length), and time-resolved changes enabled the computation of rate- and speed-based gaze metrics.

Table 1. Variables Collected for Head and Gaze Tracking.

Category	Variable	Description
Head	Head position x/y/z	The 3D position of the user's head center.
		Where the user was located in the space at each frame (left/right, up/down, forward/backward).
		The frame-to-frame positional changes, the head translation distance head movement / spatial exploration
	Head rotation x/y/z/w	The orientation/rotation of the head
		Using the difference between consecutive frame quaternions (3D rotation)
		The standard deviation of angular velocity represents head rotation variability
Gaze	Gaze origin x/y/z	The starting point of the gaze ray, where the gaze is emitted from
	Gaze direction x/y/z	The direction the gaze ray is pointing toward, represented as a 3D direction vector.
	Gaze hit position x/y/z	The spatial coordinates of where the user was looking.
		Accumulating the frame-to-frame distance between hit positions
		The gaze path length (gaze travel distance).

	The continuous changes over time, gaze movement and gaze speed metrics

2.4 Analysis

A timestamp was constructed for each sample, and the time difference between consecutive samples was computed as dt (seconds). Consecutive samples were treated as valid only when dt > 0 and when the required tracking signals were available. Large recording gaps were treated as discontinuities, and indicators were computed from valid consecutive samples within each trial.

Table 1 summarizes the four trial-level behavioral indicators derived from frame-level XR tracking data—two gaze-based metrics (gaze path rate, gaze speed variability) and two head-movement metrics (head translation path rate, head angular speed variability). Rate-based measures quantify the amount of spatial exploration per second, whereas variability measures (SD) capture instability in movement dynamics within a trial.

Four trial-level behavioral outcomes were analyzed using the pre-computed columns in the dataset. Visual exploration volume was indexed by gaze_path_rate, computed from gaze_path_length_L normalized by duration_valid_sd. Instability of gaze dynamics was indexed by gaze_speed_sd, computed from instantaneous gaze speed samples within each trial. Embodied exploration through translation was indexed by head_move_rate, computed from head_translation_length_H normalized by duration_valid_sd. Instability of head rotation dynamics was indexed by head_ang_speed_sd, computed from instantaneous head angular speed samples within each trial. For each outcome, trials with missing values in that outcome were excluded from that specific analysis.

Task-driven differences were tested by treating task difficulty as the ordered progression across the three diagnostic tasks. A task difficulty score was created from task_label so that ocular was the lowest level, cover was the middle level, and prism was the highest level. Because each participant contributed repeated trials, analyses accounted for within-person dependence. When model estimation was stable, a participant-level random-intercept model was used. When the random-effects structure was not estimable, a participant fixed-effects approach was used instead. Statistical significance was evaluated using two-sided tests with an alpha level of .05, and 95 percent confidence intervals were reported for the task-difficulty effect. All preprocessing and statistical analyses were conducted in Python using Google Colab.

3 Results

Table 2 summarizes descriptive statistics and model-based effects of task difficulty (ocular → cover → prism) on gaze and head movement indicators. Overall, increases in task difficulty were associated with greater visual exploration volume and instability, as well as greater embodied exploration through head translation.

For gaze path rate, mean values increased across. The model indicated a significant positive association with task difficulty (b = 0.32, p < .001, mixed-effects model), suggesting that gaze exploration per second increased as tasks became more difficult. For gaze speed variability, mean values also increased with task. A significant task difficulty effect was observed (b = 0.78, p = .01, participant fixed-effects), indicating greater instability in gaze dynamics during more difficult diagnostic subtasks.

For head movement rate, higher task difficulty corresponded to higher mean. The task difficulty effect was significant (b = 0.01, p < .001, participant fixed-effects), suggesting that participants engaged in greater translational head movement per second as diagnostic subtasks became more demanding. For head rotation variability, mean values showed an increasing, however, the effect did not reach statistical significance. This suggests that while head rotation dynamics tended to become more variable with task difficulty, the evidence was insufficient to conclude a reliable task-driven difference for this outcome.

Table 2. Effects of Task Difficulty on Visual and Head Exploratory Behavior.

Outcome	Mean (*SD*)			*b*	*p*	Model
	Ocular	Cover	Prism			
Gaze path rate (scan path length/s)	0.34 (0.22)	0.77 (0.52)	0.97 (0.84)	0.32	<.001	Mixed effects
Gaze speed variability (SD of gaze speed)	2.88 (1.07)	4.13(1.92)	4.33(2.65)	0.78	0.01	Fixed Effect (participant)
Head movement rate (translation/s)	0.02(0.01)	0.03(0.01)	0.03 (0.02)	0.01	<.001	Fixed Effect (participant)

Head rotation variability (SD of angular speed)	0.15(0.10)	0.22 (0.11)	0.23(0.22)	0.04	0.100	Fixed Effect (participant)

Note: *b* = unstandardized regression coefficient; *SD* = standard deviation; *p* = probability value.

4 Discussion and Conclusion

This study examined whether learners' visual and head exploration changes across three strabismus diagnosis subtasks in an XR simulation. Overall, behavioral indicators increased as tasks progressed from ocular motility to cover–uncover to prism cover, suggesting that more demanding diagnostic steps elicit greater exploration and less stable search dynamics in immersive environments. Such process level evidence complements prior XR medical education work that has largely emphasized outcomes rather than real time behavioral mechanisms. [1, 3, 5].

Consistent with prior evidence that gaze behavior shifts with exploration and visual search demands in immersive environments [11], our findings suggest that increasing diagnostic task difficulty systematically changes both visual exploration dynamics and embodied navigation in XR. These results also align with recent discussions emphasizing the value of multimodal behavioral indicators for understanding performance processes in XR-based clinical training [10]. For visual behavior, gaze path rate increased with task difficulty, indicating a larger amount of gaze travel per second. This aligns with the view that scan path-based measures reflect the volume and spatial extent of visual sampling during complex tasks [2, 12]. In addition, gaze speed variability increased, suggesting more irregular gaze dynamics when tasks became harder. Prior reviews note that gaze and pupil-based metrics can be sensitive to mental workload in simulated or gamified sensorimotor tasks, supporting the interpretation of increased variability as a marker of higher demand or reduced stability in information seeking [6].

For embodied behavior, head movement rate increased with task difficulty, suggesting greater viewpoint repositioning as diagnostic demands rose. Because visual inspection in head mounted XR is tightly coupled with head and body movement, increased translation can represent an embodied strategy to secure visibility and verify cues. This interpretation is consistent with XR research showing that cognitive/task demands can interfere with coordination and alter movement patterns in immersive settings [7–9]. By contrast, head rotation variability showed an increasing tendency but did not reach statistical significance, which may reflect constraints of the task setting or limited power given missingness in the most difficult task.

A limitation of this study is the fixed sequential order of the three diagnostic tasks. This sequence was strictly maintained to preserve the ecological validity of the simulation, as it mirrors the standard, real-world clinical workflow (i.e., observing extraocular movement, followed by identifying the deviation, and finally quantifying it). However, this fixed progression introduces potential order effects. It is possible that cumulative fatigue or cognitive depletion influenced the participants' behavioral metrics during the final and most demanding task (the prism cover test). Future research could explore balanced or counterbalanced experimental designs, where clinically feasible, to isolate the effects of task difficulty from fatigue-related order effects. Additionally, our analysis focused solely on the impact of task difficulty, without directly correlating these exploratory behaviors with the learners' actual diagnostic performance (e.g., diagnostic accuracy). While our primary goal was to emphasize process-level behavioral mechanisms over traditional post-task outcomes, comparing eye and head movement profiles between successful and unsuccessful diagnoses is highly relevant. Future research should investigate the relationship between these multimodal metrics and clinical performance scores to further validate their utility as objective assessment tools in XR training.

References

1. Bui, J., Lee, D., Murthi, S., Kantor, T., Stegink, C., Reddy, R.M.: The role of extended reality in enhancing surgical training: A narrative review. Current Challenges in Thoracic Surgery 7 (2025).
2. Fahimi, R., Bruce, N.D.: On metrics for measuring scanpath similarity. Behavior Research Methods 53(2), 609–628 (2021).
3. Siddiqui, M.F., Jabeen, S., Alwazzan, A., Vacca, S., Dalal, L., Al-Haddad, B., Kalmatov, R.: Integration of augmented reality, virtual reality, and extended reality in healthcare and medical education: A systematic review. Journal of Medical Education and Curricular Development 12 (2025).
4. Vaccaro, G.: Assessing the role of extended reality in healthcare industry: A systematic inquiry to process optimization. Master's thesis, Politecnico di Milano (2023).
5. De Witte, B., Reynaert, V., Kieken, D., Jabbour, J., Demarey, C., Dumoulin, A., Possik, J.: Immersive virtual reality learning and cognitive load: A multiple-day field study. Computers in Human Behavior 176, 108853 (2025).

6. Gorin, H., Patel, J., Qiu, Q., Merians, A., Adamovich, S., Fluet, G.: A review of the use of gaze and pupil metrics to assess mental workload in gamified and simulated sensorimotor tasks. Sensors 24(6), 1759 (2024).
7. Chua, P., Sasikumar, P., Weerasinghe, Y., Nanayakkara, S.: Motion as emotion: Detecting affect and cognitive load from free-hand gestures in VR. arXiv preprint arXiv:2409.12921 (2024).
8. Lustig, A., Wilf, M., Dudkiewicz, I., Plotnik, M.: Higher cognitive load interferes with head-hand coordination: Virtual reality-based study. Scientific Reports 13(1), 17632 (2023).
9. Poupard, M., Larrue, F., Bertrand, M., Liguoro, D., Tricot, A., Sauzéon, H.: From movement to learning: Leveraging VR behavioral metrics to evaluate cognitive load and curiosity. International Journal of Human-Computer Studies, 209 (2025).
10. González-Erena, P.V., Fernández-Guinea, S., Kourtesis, P.: Cognitive assessment and training in extended reality: Multimodal systems, clinical utility, and current challenges. Encyclopedia 5(1), 8 (2025).
11. Enders, L.R., Smith, R.J., Gordon, S.M., Ries, A.J., Touryan, J.: Gaze behavior during navigation and visual search of an open-world virtual environment. Frontiers in Psychology 12, 681042 (2021).
12. De Witte, B., Reynaert, V., Kieken, D., Jabbour, J., Demarey, C., Dumoulin, A., Possik, J.: Immersive virtual reality learning and cognitive load: A multiple-day field study. Computers in Human Behavior 176 (2025).