Divergent Patterns of Attention and Exploratory Behavior Under Increasing Task Difficulty in an XR Diagnostic

Kukhyeon Kim
Chonnam National University,
edu.kukhyeon@gmail.com

Jeeheon Ryu
Chonnam National University
jeeheon@jnu.ac.kr

ABSTRACT

This study investigates how task difficulty shapes learners' self-reported attention and multimodal exploratory behaviors in an extended reality (XR) diagnostic simulation. Twenty-one medical students completed three strabismus diagnostic subtasks of increasing difficulty (ocular motility, cover-uncover, and prism cover tests) while gaze and head movement data were continuously recorded alongside self-reported attention measured after each task. Results revealed asymmetrical divergence: self-reported attention declined significantly with task difficulty, while gaze path rate and head movement rate increased significantly. This compensation strategy indicates heightened cognitive demand. Critically, behavioral indicators carry context-dependent cognitive meanings, high exploratory activity does not uniformly signal engagement but may indicate struggle. These findings are essential for designing adaptive XR diagnostic systems that interpret behavioral signals appropriately.

Keywords

extended reality, task difficulty, attention, gaze behavior, head movement

1. INTRODUCTION

Extended Reality (XR) has gained traction as a training tool in high-stakes professional domains, offering learners a level of immersion and high contextual realism that conventional digital environments cannot replicate [1]. In clinical education, XR-based diagnostic simulations allow repeated practice of complex procedural skills in safe, controlled contexts, yet they also impose concurrent demands on visual attention, spatial navigation, and clinical reasoning that can strain learners' limited cognitive resources [2].

A key challenge in evaluating XR-based training is that most studies rely on post-task outcome measures, such as accuracy or scores, which offer limited insight into how learners allocate cognitive resources in real time. Gaze and head movement data captured continuously during task performance offer a more direct window into these moment-to-moment processes [4][5]. However, the relationship between behavioral indicators and cognitive states is not straightforward: as task demands rise, learners may simultaneously report declining attention while displaying increased exploratory effort, a behavioral compensation pattern whose meaning shifts with difficulty context [6].

The present study addresses this gap using a within-subjects design in an XR strabismus diagnostic simulation. We examine (1) how self-reported attention changes across three difficulty levels, (2) how gaze and head movement indicators vary with task difficulty, and (3) how the relationship between attention and exploratory behavior shifts depending on the difficulty context. By integrating subjective and objective behavioral measures, this study aims to provide a context-sensitive account of learner cognitive states in immersive clinical training.

2. THEORETICAL BACKROUND

2.1 XR Learning Environments and Cognitive Demands in Clinical Training

XR-based simulations have gained increasing attention in clinical education for their capacity to replicate realistic diagnostic scenarios while enabling safe, repeatable practice [2]. The Cognitive Affective Model of Immersive Learning (CAMIL) provides a theoretical account of how XR environments engage learners: technological features such as immersion and agency activate cognitive and affective processes, including cognitive load and self-regulation that mediate learning outcomes [1]. Critically, while immersion enhances engagement, it can simultaneously impose extraneous cognitive demands through the need to navigate three-dimensional space and coordinate visual and physical actions alongside core task performance. This dual nature makes it particularly important to characterize how learners' cognitive states change as task demands escalate within XR simulations.

2.2 Task Difficulty, Attention, and Behavioral Compensation

According to Cognitive Load Theory, human working memory is limited, and rising task complexity increases intrinsic load, progressively depleting the cognitive resources available for sustained attention [3]. As task demands increase, learners are expected to report decreased attentional focus, which can be reliably measured using validated subjective mental effort scales [7]. However, the relationship between subjective attention and observable behavior is not straightforward. Rather than reducing exploratory effort in

Kukhyeon Kim, and Jeeheon Ryu. Divergent Patterns of Attention and Exploratory Behavior Under Increasing Task Difficulty in an XR Diagnostic. In Anthony Botelho, Maria Mercedes T. Rodrigo, Adish Singla, Hiroaki Ogata, Hyojeong So, and Young Hoan Cho (eds.) Proceedings of the 19th International Conference on Educational Data Mining, Seoul, Republic of Korea, June, 2026, pp. 637–641. International Educational Data Mining Society (2026).

https://doi.org/10.5281/zenodo.21039802

proportion to attentional decline, learners may instead intensify their visual and physical search to compensate for depleted cognitive resources, a pattern termed behavioral compensation [4]. Whether this compensatory pattern emerges, and how its meaning shifts with difficulty context, remains an underexplored question. Crucially, behavioral compensation is conceptualized here as a strategic redirection of behavioral effort when attentional resources are depleted, distinguishing it from random or fatigue-driven movement: learners actively intensify exploration to gather diagnostic cues despite reduced subjective attention [10, 11]. This framework predicts not only main effects of difficulty on both attention and behavior, but also a context-dependent reversal of their relationship at high difficulty, where preserved attentional capacity becomes increasingly coupled with strategic exploratory effort [12].

2.3 Multimodal Behavioral Indicators: Gaze and Head Movement

Eye gaze and head movement provide continuous, non-intrusive indices of cognitive dynamics during task performance that retrospective self-reports cannot capture [5]. Gaze path rate reflects the volume and spatial extent of visual exploration, while gaze speed variability indexes instability in information-seeking dynamics. In XR specifically, head translation path rate captures embodied spatial exploration, as learners must physically reposition themselves to acquire visual information within three-dimensional space [8]. Critically, the functional meaning of these indicators may vary depending on difficulty context: the same level of gaze exploration may reflect unfocused, scattered search at low difficulty, but strategic, resource-dependent search at high difficulty [6]. Integrating subjective attention with multimodal behavioral data is therefore essential for accurate cognitive assessment.

.

3. METHOD

3.1 Participants

Twenty-one medical students (13 males, 8 females; third- or fourth year) participated in this study. All had completed relevant ophthalmology coursework and practical training in strabismus prior to participation. The study was approved by the Institutional Review Board (No. 1040198-221021-HR-128-02).

3.2 Procedure and Task Design

Participants completed three strabismus diagnostic subtasks using a HoloLens 2 head mounted display, following a 10-minute familiarization session and eye-tracking calibration. The tasks were administered in a fixed order mirroring standard clinical workflow with progressive cognitive demands: (1) ocular motility test (Level 1; low difficulty), in which participants positioned themselves approximately 40 cm from the virtual patient and moved a fixation stick along an H-pattern while tracking the patient's eye movements to determine the presence of strabismus; (2) cover-uncover test (Level 2; medium difficulty), in which participants alternately covered each of the virtual patient's eyes with an occluder and observed refixation movements to identify both the presence and the type of strabismus; and (3) prism cover test (Level 3; high difficulty), in which participants placed a prism over the non-deviating eye while applying the occluder to the strabismic eye, then iteratively adjusted the prism degree until eye movement was no longer observed in order to quantify the angle of deviation. The progressive difficulty structure reflects escalating demands in visual search, pattern recognition, and procedural reasoning beyond passive vision screening, and is consistent with established clinical training standards and was validated by ophthalmology specialists. Trial duration was limited to three minutes for Levels 1 and 2, and five minutes for Level 3. A within-subjects design was employed, with each participant completing all three tasks.

3.3 Measures

Self-reported attention was assessed immediately after each task using a 7-point Likert scale [9], where higher scores indicate greater attentional focus. Four behavioral indicators were derived from continuous XR tracking logs. *Gaze path rate* (GPR) and *gaze speed variability* (SD of instantaneous gaze speed) served as indices of visual exploration volume and stability, respectively, computed from frame-to-frame gaze hit position data. *Head movement rate* (HMR) and *head angular speed variability* (SD of angular speed) served as indices of embodied exploration, derived from head position and rotation quaternion data sampled at each frame.

3.4 Analysis

Task difficulty was treated as an ordered three-level predictor. To address RQ1 and RQ2, within-person dependence was accounted for using participant-level random-intercept linear mixed models (LMM), a participant fixed-effects approach was substituted when random-effects structures were not estimable. To address RQ3, Pearson correlations between self-reported attention and each behavioral metric were computed separately at each difficulty level to examine context dependent shifts in the attention and exploration relationship. All analyses were conducted in Python (Google Colab), with $\alpha = .05$.

Note on the dataset. A complementary descriptive analysis of the same XR strabismus simulation dataset has been reported in the proceedings of iLRN 2026 [13]. That prior work focused exclusively on comparing behavioral indicators across difficulty levels. The present study extends this work by introducing self-reported attention as a subjective cognitive measure, examining the difficulty-contingent relationship between attention and exploratory behavior, and interpreting findings through Cognitive Load Theory to propose a behavioral compensation mechanism.

4. RESULTS

Table 1 summarizes descriptive statistics and model-based effects for all five outcome variables across the three difficulty levels.

4.1 RQ1. Effects of Task Difficulty on Self-Reported Attention

Self-reported attention decreased significantly as task difficulty increased ($\beta = -0.184$, $p = .008$, LMM). Mean scores declined monotonically from 6.20 (*SD* = 0.62) at Level 1 to 6.06 (*SD* = 0.74) at Level 2 and 5.83 (*SD* = 0.54) at Level 3, confirming that learners experienced progressively greater cognitive strain across the three diagnostic subtasks.

4.2 RQ2. Effects of Task Difficulty on Exploratory Behavior

Increased task difficulty was associated with significant increases across three of the four behavioral indicators. Gaze path rate rose approximately 2.85-fold from Level 1 to Level 3 ($\beta = 0.32, p < .001$, LMM), indicating that learners engaged in substantially broader visual scanning under higher demand. Gaze speed variability also increased significantly ($\beta = 0.78$, $p = .010$, fixed effects), reflecting greater instability in visual search dynamics. Head movement rate similarly increased with difficulty ($\beta = 0.01, p < .001$, fixed effects), suggesting that learners engaged in greater physical repositioning as diagnostic demands escalated. Head angular speed variability showed an increasing trend but did not reach statistical significance (β = 0.04, p = .100). Together, these results indicate that despite declining subjective attention, learners intensified their exploratory effort across difficulty levels, a pattern consistent with behavioral compensation under increasing cognitive load. Notably, while the decline in self-reported attention was modest (5.9% reduction across levels), behavioral indicators showed substantially larger relative increases (gaze path rate: 185%; gaze speed variability: 50%; head movement rate: 50%), suggesting that objective behavioral metrics may capture cognitive demand shifts more sensitively than subjective reports.

4.3 RQ3. Difficulty-Contingent Shifts in the Attention–Exploration Relationship

The relationship between self-reported attention and GPR shifted in direction across difficulty levels (Table 2). At Levels 1 and 2, negative correlations were observed (r = -0.217 and r = -0.176, respectively), indicating that learners reporting lower attention exhibited broader visual exploration patterns, consistent with unfocused, scattered search. In contrast, at Level 3, the correlation reversed to positive (r = 0.342), despite a statistically significant decline in mean attention across the cohort (L1: M = 6.20 → L3: M = 5.83, β = -0.18, p = .008). This reversal is notable: the substantial increase in mean GPR (from 0.34 at L1 to 0.97 at L3) coincided with decreased attention, suggesting that visual exploration intensity increased not because of greater attentional resources, but possibly due to heightened task demand. Within this constrained attentional environment, the positive correlation (r = 0.342) indicates that learners who retained relatively higher attention tended to exhibit more extensive gaze-based exploration, a pattern consistent with a compensatory mechanism whereby preserved attentional capacity is redirected toward broader information gathering.

For HMR, negative correlations persisted across all three levels (Level 1: r = -0.182; Level 2: r = -0.300; Level 3: r = -0.055), with magnitude substantially attenuated at Level 3. This progressive dissociation between attention and head-based exploration under increasing demand suggests that gaze became the primary behavioral marker of cognitive resource allocation at peak difficulty, whereas head movement decoupled from attentional availability, possibly reflecting a shift toward more efficient, gaze-centered diagnostic strategies.

Table 1. Descriptive Statistics and Model Effects by Task Difficulty

Outcome	L1 M (*SD*)	L2 M (*SD*)	L3 M (*SD*)	β	p	Model
Attention	6.20 (0.62)	6.06 (0.74)	5.83 (0.54)	−0.18	.008	LMM
Gaze path rate	0.34 (0.22)	0.77 (0.52)	0.97 (0.84)	0.32	<.001	LMM
Gaze speed variability	2.88 (1.07)	4.13 (1.92)	4.33 (2.65)	0.78	.010	Fixed
Head movement rate	0.02 (0.01)	0.03 (0.01)	0.03 (0.02)	0.01	<.001	Fixed
Head rotation variability	0.15 (0.10)	0.22 (0.11)	0.23 (0.22)	0.04	.100	Fixed

Note. L1 = ocular motility; L2 = cover–uncover; L3 = prism cover. β = unstandardized coefficient. LMM = linear mixed model. Fixed = participant fixed effects.

Table 2. Pearson Correlations Between Attention and Exploratory Metrics by Difficulty Level

	Level 1	Level 2	Level 3
Attention – GPR	−0.217	−0.176	+0.342
Attention – HMR	−0.182	−0.300	−0.055

Note. GPR = gaze path rate; HMR = head movement rate. Bold indicates direction reversal.

5. DISCUSSION AND CONCLUSION

This study examined how task difficulty shapes self-reported attention and multimodal exploratory behaviors in an XR diagnostic simulation, and how the relationship between these measures shifts with difficulty context.

5.1 Declining Attention and Behavioral Compensation

Consistent with Cognitive Load Theory [3], self-reported attention declined monotonically as diagnostic task demands escalated. Concurrently, gaze path rate and head movement rate increased significantly, presenting an asymmetrical pattern in which subjective attentional resources were depleted while behavioral exploratory effort was amplified. This divergence is interpretable as a behavioral compensation strategy: despite declining attentional capacity, learners persistently expanded their visual search for diagnostic cues. The finding underscores that behavioral quantity (gaze/head movement) and attentional quality (self-reported focus) do not move in tandem, and that exploration metrics alone may misrepresent cognitive engagement without accounting for difficulty context and concurrent attentional constraints. An alternative interpretation, that increased eye movements at higher difficulty reflect visual strain rather than cognitive compensation, is less consistent with the observed selective patterns: gaze path rate increased significantly while head rotation variability did not, and the GPR–attention correlation reversed at Level 3 — patterns not predicted by a fatigue-based account[11].

5.2 Difficulty-Contingent Reversal of the Attention–Exploration Relationship

The most theoretically significant finding is the reversal of the attention−GPR correlation at Level 3. At lower difficulty levels, lower attention was associated with broader exploration, consistent with unfocused, scattered search. At the highest difficulty level, this relationship reversed within a substantially constrained attentional environment. Within this depleted-attention context, learners who retained relatively higher attentional resources tended to exhibit more extensive gaze-based exploration.

This pattern suggests a compensatory mechanism: when task difficulty depletes attentional capacity across the cohort, learners with preserved (relative) resources redirect their search effort toward more thorough visual exploration to compensate for degraded cognitive focus. The reversal does not indicate that higher absolute attention predicts exploration; rather, it reflects how the same cognitive resource (attention) is allocated differently under distinct difficulty contexts.

This finding aligns with the CAMIL model's prediction that high cognitive demands amplify performance heterogeneity [1] and underscores the critical risk of misclassifying learner cognitive states when behavioral indicators are interpreted without considering concurrent difficulty level and attentional availability [5].

5.3 Implications, Limitations, and Conclusion

These findings carry direct implications for adaptive XR learning system design. Because the same behavioral signal carries different cognitive meanings across difficulty levels and attentional states, real-time learner state monitoring requires multimodal data that combines objective behavioral indicators (gaze, head movement) with subjective attention measures and explicit difficulty-level context. A learner exhibiting simultaneously low attention and high gaze exploration under peak demand may not be "strategic" but rather "compensatory"—attempting to maintain performance through expanded search despite depleted attentional capacity. Conversely, low exploration coupled with low attention under high demand may indicate cognitive overload and risk of disengagement, warranting targeted adaptive scaffolding.

Several limitations should be noted. The small sample size limits generalizability, and the modest strength of the attention–exploration relationship suggests substantial individual variability. The fixed task order, while ecologically valid, introduces potential order and fatigue effects. Future research should employ larger, multisite samples to confirm the replicability of the difficulty-contingent reversal, examine how these behavioral-attentional profiles relate to diagnostic accuracy outcomes, and investigate whether real-time multimodal monitoring can support adaptive difficulty adjustment in clinical XR training.

This study has several limitations. The modest sample size (n = 21) limits generalizability, though the consistent direction of effects across multiple indicators warrants larger-scale replication. The fixed task order, retained for ecological validity in line with standard clinical workflow, cannot fully separate difficulty effects from order or fatigue effects; counterbalanced designs are needed where clinically feasible. Self-reported attention, collected post-task, may not capture moment-to-moment fluctuations; future research should incorporate diverse cognitive load measures (e.g., pupillometry, EEG) to formally test the behavioral compensation hypothesis. Finally, future work should examine how these multimodal behavioral signatures predict diagnostic accuracy to validate their utility as adaptive assessment tools in XR training

In conclusion, this study provides empirical evidence that task difficulty produces divergent, difficulty-dependent trajectories in attention and exploratory behavior in XR. Critically, the functional meaning of behavioral indicators (gaze, head movement) is not static but context-dependent, shaped by both difficulty level and the learner's residual attentional capacity. These findings advance the theoretical understanding of learner cognitive states in immersive environments and offer a nuanced behavioral basis for designing more responsive, context-aware clinical training systems that account for the complex interplay between objective exploration and subjective attentional availability.

6. REFERENCES

[1] Makransky, G. and Petersen, G. B. 2021. The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review* 33, 937–958. DOI= https://doi.org/10.1007/s10648-020-09586-2

[2] Laspro, M., Groysman, L., Verzella, A. N., Kimberly, L. L., and Flores, R. L. 2023. The use of virtual reality in surgical training: Implications for education, patient safety, and global health equity. *Surgeries* 4, 4, 635–646. DOI= https://doi.org/10.3390/surgeries4040056

[3] Sweller, J. 2019. Cognitive load theory and educational technology. *Educational Technology Research and Development* 68, 1, 1–16. DOI= https://doi.org/10.1007/s11423-019-09701-3

[4] Gorin, H., Patel, J., Qiu, Q., Merians, A., Adamovich, S., and Fluet, G. 2024. A review of the use of gaze and pupil metrics to assess mental workload in gamified and simulated sensorimotor tasks. *Sensors* 24, 6, 1759. DOI= https://doi.org/10.3390/s24061759

[5] Mu, S., Cui, M., and Huang, X. 2020. Multimodal data fusion in learning analytics: A systematic review. *Sensors* 20, 23, 6856. DOI= https://doi.org/10.3390/s20236856

[6] Schirm, J., Gómez-Vargas, A. R., Perusquía-Hernández, M., Skarbez, R. T., Isoyama, N., Uchiyama, H., and Kiyokawa, K. 2023. Identification of language-induced mental load from eye behaviors in virtual reality. *Sensors* 23, 15, 6667. DOI= https://doi.org/10.3390/s23156667

[7] Paas, F. G. W. C., van Merriënboer, J. J. G., and Adam, J. J. 1994. Measurement of cognitive load in instructional research. *Perceptual and Motor Skills* 79, 1, 419–430. DOI= https://doi.org/10.2466/pms.1994.79.1.419

[8] Lustig, A., Wilf, M., Dudkiewicz, I., and Plotnik, M. 2023. Higher cognitive load interferes with head-hand coordination: Virtual reality-based study. *Scientific Reports* 13., 17632. DOI= https://doi.org/10.1038/s41598-023-43337-x

[9] Georgiou, Y., & Kyza, E. A. 2017. The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human-Computer Studies, 98*, 24-37. https://doi.org/10.1016/j.ijhcs.2016.09.014

[10] Martinez-Cedillo, A. P., Gavrila, N., Mishra, A., Geangu, E., and Foulsham, T. 2025. Cognitive load affects gaze dynamics during real-world tasks. *Experimental Brain Research* 243, 4, 82. DOI= https://doi.org/10.1007/s00221-025-07037-4

[11] Walter, K., and Bex, P. 2021. Cognitive load influences oculomotor behavior in natural scenes. *Scientific Reports* 11, 12405. DOI= https://doi.org/10.1038/s41598-021-91845-5

[12] Nasri, M., Kosa, M., Chukoskie, L., Moghaddam, M., and Harteveld, C. 2024. Exploring eye tracking to detect cognitive load in complex virtual reality training. In *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 51–54. DOI= https://doi.org/10.1109/ISMAR-Adjunct64951.2024.00022

[13] Kim, K., Han, D., Yang, E., and Ryu, J. 2026. Analysis of Attention Based on Gaze and Head Movements under Task Difficulty in an Extended Reality Diagnostic Simulation. In