Embodied sampling: An augmented reality investigation of visual attention

June 2025

Wilbert Tabone[1], Joost de Winter[1], Renchi Zhang[1], Christopher D. Wickens[3], Mark van de Ruit[2], Yke Bauke Eisma[1]

[1] Department of Cognitive Robotics, Faculty of Mechanical Engineering, Delft University of Technology
[2] Department of BioMechanical Engineering, Faculty of Mechanical Engineering, Delft University of Technology
[3] Department of Psychology, Colorado State University, Fort Collins, CO, USA

Abstract

Traditional visual sampling research often uses static setups, which limits the understanding of attention in dynamic environments. This study used augmented reality (AR) to investigate how display characteristics and physical effort (eye, head, or torso rotation, and ambulation) affect visual attention, task performance, and physical movement during a multi-dial monitoring task. Results showed that the highest performance (64.0% detections) was achieved when high-bandwidth information was centrally presented with maximal physical constraints. In contrast, performance was considerably lower (between 33.9% and 48.2%) when information was less accessible or when participants had greater freedom of movement with larger or spatially distributed displays. In summary, although participants increased head rotation and body movement as a coping mechanism for larger displays, this compensatory activity did not lead to more effective attention allocation. These findings suggest that interfaces in AR and control room settings should centralize vital information to minimize physical demands.

Keywords: Embodied interaction, head-locked vs. world-locked displays, distributed attention, eye-tracking, eye-, head-, and body movements

Introduction

The ability of humans to sample information from multiple sources is a key element of safe and efficient performance in operational environments. Seminal work by Fitts et al. (1950) provided early insights into visual attention distribution by examining pilots' eye movements during instrument landing, and found that fixation frequency was indicative of instrument importance. Building upon this, Senders (1964, 1983) introduced a theory-driven approach to visual sampling by proposing that humans sample instruments at a rate proportional to the bandwidth of the signal displayed, akin to the Nyquist-Shannon sampling theorem. This approach to information seeking resembles theories of information foraging, which posit that humans adapt their information-gathering strategies to maximize the rate of gain of information relative to the costs of searching (Pirolli & Card, 1999).

Senders' paradigm, which involved participants monitoring a bank of dials with different signal bandwidths, has been foundational in human factors, and human-computer interaction (Hancock et al., 2019; Rouse, 1981). His research demonstrated a linear relationship between signal bandwidth and visual sampling measures such as glance rate and dwell time. More recently, Eisma et al. (2018, 2024) replicated and extended Senders' (1983) work using modern eye-tracking equipment and larger participant samples. Their findings corroborated Senders' original results regarding the impact of signal bandwidth on sampling behaviour. Additionally, Eisma et al. provided evidence that participants act as conditional samplers by adjusting their gaze based on the momentary state (e.g., pointer angle and velocity, or 'saliency') of the dials, and they investigated the effect of effort operationalized as how the dials were positioned on the screen. Their results were found to be consistent with the SEEV model, a computational attention framework which integrates top-down expectancy (i.e., signal bandwidth), bottom-up saliency, and the effort required to shift gaze (Wickens, 2008).

A common characteristic of visual sampling studies, including those by Senders and more recent replications (e.g., Eisma et al., 2018, 2024; Gasse et al., 2025), is the use of static setups. Participants were seated with their heads stabilized by a chinrest to ensure accurate eye-tracking and maintain a fixed distance and orientation to the dials. While this provides high experimental control, it does not necessarily capture the dynamics of attention allocation in real-world environments. For example, operators in large-scale industrial control rooms often need to monitor displays distributed across multiple panels. This does not just involve eye movements, but also head and full-body reorientations to bring information sources into their field of view (Kovesdi et al., 2018; Le Blanc et al., 2015).

Augmented reality (AR) provides the capability to present spatially distributed information while enabling measurement of visual sampling behaviour using eye-tracking (e.g., Sidenmark & Gellersen, 2019). The current study used AR to investigate how display spatial configuration and interactivity affect visual sampling processes, task performance, and perceived effort. Our approach follows a model by Warden et al. (2024) which says that the effort required to access information is defined by a scale based on which muscle groups are successively activated as the visual angle of separation increases. This scale begins with the 'eye field', where eye movements are adequate to access information, extending to 20–25° of eccentricity. Beyond that angular distance, one enters the 'head field', in which head rotation becomes necessary, and finally the 'body field' which is used at the greatest separations and requires torso rotation. While Warden et al. found that increasing visual separation resulted in increased head movements to preserve accuracy without major time or accuracy penalties, more recent research by Poole et al. (2025) using an AR device with a spatial integration task found that head movements became costly to response time when information entered the head field.

Building on earlier studies that engaged the 'head field', the present study was designed to investigate the full hierarchy of this motor-based effort scale. We varied the physical constraints on information access, ranging from highly constrained head-locked systems (primarily engaging eye movements within the 'eye field') to world-locked screens viewed by seated participants (requiring head rotation and engaging the 'head field'), and finally to relatively unconstrained world-locked screens viewed by standing participants, intended to evoke torso rotation and allow a degree ambulation, thus engaging the 'body field'.

In this study, manipulated the display setup along an ordinal scale of information access costs, ranging from head-locked screens (engaging only eye movements) to world-locked screens that required head rotation, and finally to highly distributed layouts encouraging full-body reorientations and ambulation. We assessed how these motor behaviors, in conjunction with variations in screen type and spatial arrangement (e.g., screen size, single wall, two walls) and task-intrinsic visual effort (manipulated by dial bandwidth distribution), affected attention allocation and task performance in detecting threshold crossings. We hypothesized that increasing embodied interaction freedom and information distribution, particularly when coupled with high task-intrinsic visual effort (i.e., fast-moving dials in less central locations), will decrease overall task performance and attention allocation efficiency, while leading to a compensatory increase in head and body movements.

Methods

Participants and recruitment

A total of 30 participants (23 male, 7 female), aged between 21 and 35 (M = 27.33, SD = 3.93), were recruited. The 30 participants were of nine different nationalities, namely Cypriot (1), Chinese (8), Dutch (10), German (1), Indian (1), Italian (3), Maltese (3), South African (2), and Russian (1). From the participant pool, 80% (n = 24) indicated that they had used a VR headset before, while 63% (n = 19) had previously used an AR application or played AR games such as Pokémon GO.

Recruitment was opportunistic, with most participants recruited from the offices adjacent to the experiment room, other university faculties, and people known to the experimenters. Participants could book a time slot using an online calendar system. The experiment ran for 3 weeks between March and April 2025, and lasted approximately 45 minutes per participant. No remuneration was offered.

The criteria for participating in the experiment were indicated in the calendar system and in a subsequent confirmation email. Specifically, participants were only eligible to participate if they were over 18 years of age, had a good command of English, did not suffer from severe mobility issues, and did not suffer from epilepsy, claustrophobia, or feelings of disorientation. Moreover, we asked participants to wear lenses if they had prescription glasses, and if they had long hair, to tie it back on the day of the experiment.

The experiment was approved by the Human Research Ethics Committee (HREC) of the TU Delft under reference number 5106.

Apparatus and resources

This study used the Magic Leap 2, an AR device that can provide a digital overlay of 45° H × 54° V, at a resolution of 1440 × 1760 pixels per eye. The Magic Leap generates AR content by reflecting light from a Liquid Crystal on Silicon (LCoS) panel into a series of three color-selective waveguides that direct the light to the user's eye (Magic Leap, 2022). The AR application was built in Unity (v2022.3.38f1) using the Mixed Reality Toolkit (MRTK3). The application was based on the OpenXR framework (v1.11.0) and used the Magic Leap SDK and the OpenXR Plugin for Unity. The application for the experiment was developed and compiled using an Alienware computer with Intel(R) Core(™) i9-9900K CPU @ 3.60 GHz, 64 GB RAM, NVIDIA GeForce RTX 2080 Ti with 11 GB of memory, and 3.68 TB storage space, and was run on the Magic Leap's onboard computer.

The dial videos were taken from the supplementary material of a previous experiment (Eisma et al., 2018), which was a replication study of early work by Senders (1983). The videos showed a total of six dials, with pointers that moved at different speeds (signal bandwidths: 0.03, 0.05, 0.12, 0.20, 0.32, and 0.48 Hz). For the present experiment, only dial configurations of the lowest effort level (level 1, with the fastest-moving dials placed centrally) and the highest effort level (level 7, with the fastest-moving dials placed in the opposite corners) were used.

Experimental conditions

Ten experimental conditions were created (Table 1). The conditions differed in anchoring: either head-locked, i.e., fixed screen coordinates of the AR headset's display, or world-locked, where AR elements remained anchored to a fixed position in world space. Furthermore, participants were either seated and restricted to head movements, or free to stand and given the opportunity to walk.

In the head-locked conditions (A and B), seated participants experienced the dial screen fixed directly before their eyes. Although participants could move their heads, the experience of conditions A and B may have been comparable to studies using a head support or chinrest (Eisma et al., 2018; Senders, 1964), in that the screen remained directly in front of them throughout the trial. The depicted screen measured 1.49 m wide by 0.96 m high and was positioned at a distance of 2.30 m, resulting in a horizontal visual field spanning 36° and a vertical visual field spanning 24°.

The remaining conditions employed a world-locked screen. In conditions C and D, seated participants viewed a screen simulating a standard desktop monitor. The screen was 1.93 m wide by 1.04 m high and was positioned at a distance of 2.23 m, resulting in a horizontal visual field spanning 47° and a vertical visual field spanning 26°.

In Conditions E and F, standing participants were positioned to enable full body movement before a large screen of dials, akin to a control room meta-display. The screen was 6.04 m wide by 3.24 m high, and was positioned at a distance of 3.14 m. This resulted in a horizontal and vertical visual field spanning 88° and 55°, respectively, when standing in front of its center. Since the AR device could only cover 45° horizontally, participants had to rotate their heads to sample all dials.

In conditions G and H, participants were instructed that they could walk to observe dials arranged as a single bank along one large wall, viewed from a distance. The bank of dials was 31.5 m long and 4.95 m high, with its center positioned 13.1 m from the participant's starting position. The screen was slightly turned (8°) towards the participant, for a better view of all six dials. From the participant's starting point, the dial screen subtended approximately 106° horizontally.

In conditions I and J, participants observed two screens on opposite walls. The left and right banks of dials were 16.7 m long and 4.36 m high from the participant's starting position. Both screens were turned 8° inwards, i.e., towards the participant. From the participant's starting point, the dial screen spanned approximately 50° horizontally.

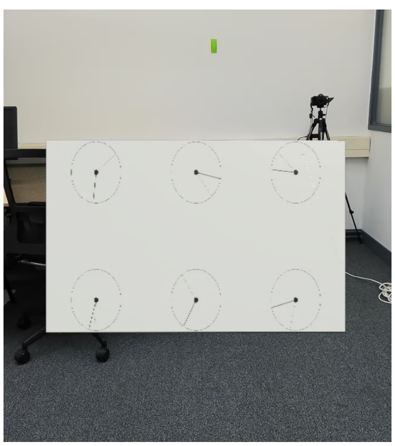
The experiment took place in a room 6.65 m long and 3.21 m wide, with the participant positioned on a chair (conditions A–D) or standing (conditions E–J) in the middle of that room, oriented along its long axis. A general factor for all conditions was that the device remained tethered to a computer via a 3 m data cable for video capture, requiring participants to stop walking if the cable became taut. Due to this cable and other constraints, such as the experimenter's desk which was located behind the participant, the participant in conditions E–J was physically limited to walking approximately 2 m forward and 1 m backward. This physical limitation meant that while full head and body rotation was possible, extensive ambulation, as a complete compensatory strategy for widely distributed information, was not feasible for participants.

Table 1.
Experimental conditions

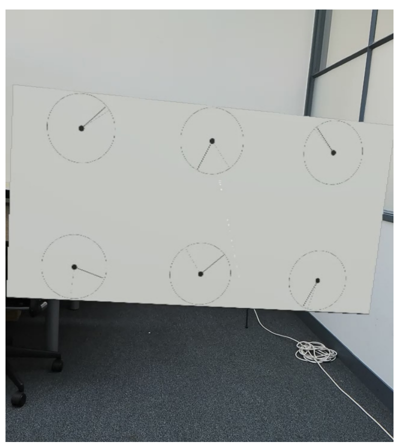
Condition	**Screen anchoring**	**Participants' posture**	**Participants' freedom of movement**	**Screen type and dial configuration**	**Dial configuration effort level**
A	Head	Sitting	Only head	Screen, 3×2	Low
B	Head	Sitting	Only head	Screen, 3×2	High
C	World	Sitting	Only head	Screen, 3×2	Low
D	World	Sitting	Only head	Screen, 3×2	High
E	World	Standing	Entire body	Large screen, 3×2	Low
F	World	Standing	Entire body	Large screen, 3×2	High
G	World	Standing	Entire body	Single wall, 6×1	Low
H	World	Standing	Entire body	Single wall, 6×1	High
I	World	Standing	Entire body	Two walls, 3×1	Low
J	World	Standing	Entire body	Two walls, 3×1	High

Note. In conditions I and J, 3 dials (corresponding to the top row of the 3x2 bank of dials) were positioned on the left side of the room, and 3 dials (corresponding to the bottom row of the 3x2 bank of dials) were positioned on the right side of the room.

Figure 1 provides an overview of the screens in the environment.



Screenshot of video capture of the Magic Leap 2 as it could have appeared in condition A and B. The screen with the dials is fixed in the middle of the participant's field of view, i.e., a head-locked presentation.



Screenshot of video capture of the Magic Leap 2 as it could have appeared in condition C and D. At the time of the screenshot, the participant had their head slightly turned to reduce eye movements toward a dial.

Screenshot of video capture of the Magic Leap 2 as it could have appeared in condition E and F. In this case, the participant is standing, and has taken a step back to be able to see more of the screen at once, which is three times as large as in conditions A–D.

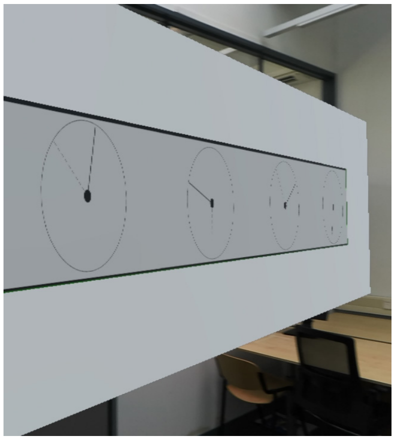
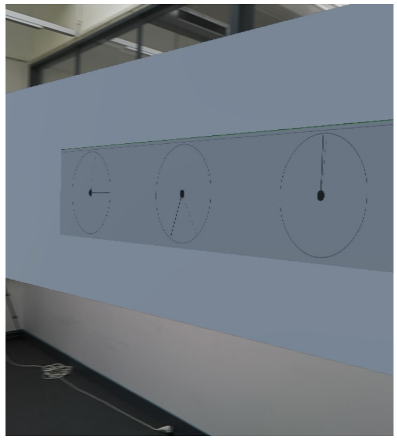
 Screenshot of video capture of the Magic Leap 2 as it could have appeared in condition G and H. In this case, the participant took a step forward to gain a better view of four of the six dials presented on the left side of the room.	 Screenshot of video capture of the Magic Leap 2 as it could have appeared in condition I and J. Here, the participant is looking at the three dials on the right side of the room. The other three dials are presented on the left side of the room.	

Figure 1. Screenshots of screen captures of the conditions of the experiment

Experiment procedure

The participants were welcomed and invited to sign the consent form, which also contained instructions about the procedure. Following signing, the experimenter reiterated the instructions. The participants were told that their goal was to press the bumper button of the controller whenever any of the dials' needles passed through a threshold (a dotted line on the dial) in any direction. Each trial had a sitting or standing condition. It was told that in the latter case, dials may appear anywhere in the room except behind them; moreover, they were free to walk in this condition, and also take a step back if necessary. Finally, participants were informed that a question panel would be shown at the end of each trial, where they would rate their eye-movement effort and subsequently elaborate on the strategies, in terms of eye, head, and body movement that they used to accomplish the task.

Following the instructions, participants were invited to sit in the starting position, after which the device (cleaned between each participant) and its computer pack were worn, and the controller was held in the participant's dominant hand. A fit check and eye calibration process was conducted using the 'Custom Fit' in-built application onboard the Magic Leap. The fit check ensured that the device was secured around the participant's head and that they could clearly see everything on the screen. Eye calibration followed, where participants were instructed to gaze at a moving point on the screen until the process was complete.

Once the calibration process was complete, participants were asked to open a demo application that contained the trials presented in sequential order and with a runtime of just 5 seconds. The goal was to demonstrate how the experiment would flow. Once the participant indicated that they were confident enough to start the experiment, participants closed the trial application, and the experimenter launched the actual application.

The participants underwent 10 trials of 90 seconds each, which were presented in a counterbalanced order that differed for each participant. At the start of each trial, a panel was shown, instructing participants to be at the starting position. After 5 seconds, the panel vanished, and the 90-second video of the corresponding trial condition would be shown. At the end of each trial, a question panel was displayed, and participants had to verbally report a number between 1 (lowest) and 10 (highest) for eye-movement effort. The second question relating to the strategies used to accomplish the task was open-ended and recorded using a Philips VoiceTracer stereo voice recorder. Once the short interview was conducted, participants were told that they could move to the subsequent trial, at which point, they pressed the cyan button and waited. During this time, the experimenter indicated whether the following trial would be a 'sitting' or 'standing' trial and made the necessary arrangements by either placing or removing the chair.

Once the 5th trial was finished, participants were informed that they were halfway through the experiment, and an opportunity to take a break was offered, but none of the participants opted for it. At the end of the experiment, the device was placed on the chair, and participants were asked to fill in the demographic questionnaire. Finally, participants were thanked and offered the opportunity to ask any questions they may have had about the experiment's goals.

Data processing and dependent variables

Data was recorded in CSV files at 60 Hz during the experiment. The log files included the gaze *x*, *y*, and *z* coordinates relative to the display depicting the dials, the *x*, *y*, and *z* coordinates of the Magic Leap device in world space, and a boolean indicating whether the bumper button was pressed at that time. Due to a programming omission, a quaternion representing the device's orientation was recorded for only 19 of the 30 participants.

For each participant and trial, raw CSV files were read into a unified MATLAB data structure. The time-series data was temporally aligned by resampling it from its original 60 Hz logging rate to a uniform 100 Hz using linear interpolation for continuous data and a 'previous value' method for discrete events. Periods identified by the eye tracker as blinks or eye closures were marked as invalid, and their corresponding gaze coordinates on the dial screen were set to NaN to ensure they were excluded from the calculation of attention-based measures.

The following measures were calculated from the data, for each participant and trial:

1. *Performance score (%):* The percentage of threshold crossings for which the participant correctly pressed the controller's bumper button. A response was counted as correct if the button press occurred within a 0.5-second (50-sample) window relative to the ground-truth time of a pointer crossing.
2. *Attention on dial (% of 90 s trial time):* The percentage of the total trial time for which the participant's valid gaze coordinates fell within the predefined area of interest (AOI) for each of the six dials. This allocation was determined by classifying the gaze coordinates into one of six rectangular regions defined by coordinate boundaries. These boundaries were adjusted for each experimental condition to match the spatial layout of the dials.
3. *Slope of attention on dial vs. bandwidth (%/Hz):* The slope of a least-squares linear regression fit between the percentage of attention on each dial (Measure 2) and the dial's known signal bandwidth (n = 6 dials). A steeper slope indicates that attention was more effectively distributed, with higher bandwidth dials that change more frequently receiving a proportionally greater amount of visual attention. It is important to note that the goodness-of-fit for individual participants and conditions varied, and the relationship between attention and bandwidth was not always strongly linear across all dials. Hence, the slope is interpreted as an indicator of the *tendency* for attention to match bandwidth.

4. *Self-reported eye-movement effort (1–10 scale):* Participants provided a rating of the perceived effort of their eye movements on a scale from 1 (very low) to 10 (very high). This verbal rating was recorded by the experimenter immediately after completing each 90-second trial.
5. *Total head orientation change (°):* The cumulative angular change in head orientation throughout a trial, representing an objective measure of head movement. This was calculated by summing the angular difference between consecutive head orientation quaternions.
6. *Total distance traveled (m):* The total distance the participant's head moved through the physical laboratory space, representing the extent of ambulatory activity. This measure also served as an indicator of participants' inclination to walk given the experimental constraints.

The eye-tracking measures (variables 2 and 3) were only calculated for trials where less than 20% of the data was classified as invalid. In total, this accounted for 76.3% of the 300 trials.

Finally, we created a post-trial strategy report consisting of summaries of the monitoring strategies described by participants in short, open-ended interviews after each trial. The 300 recordings were automatically transcribed using the AssemblyAI (2025) API, with the multi-speaker detection feature enabled. The summaries were grouped by trial condition into text files. Next, the 10 combined text files were then uploaded to Gemini 2.5 for contextual summarisation that followed a method used in a previous study that validated a similar approach (Tabone & De Winter, 2023). Details can be found in Appendix A.

Statistical Analyses

The statistical analysis consisted of descriptive statistics, including means, standard deviations, and 95% confidence intervals for within-subject designs (Morey, 2008) or for independent-samples in case of unequal variances. Because the experimental stimuli differed substantially from each other, the effects are often highly statistically significant. This significance can be inferred from the non-overlapping confidence intervals shown in the figures (Cumming, 2009). Therefore, our primary interest was in describing these effects, with limited use of p-values. In a number of cases, we performed paired-samples t-tests to compare two conditions, using an alpha level of 0.005 (Benjamin et al., 2018).

Results

From the analysis of the timings of the button presses in relation to the ground-truth pointer threshold crossings (see Figure 2), it can be seen that performance (i.e., the percentage of detected threshold crossings) was highest (64.0%) for condition A, which is the condition where the fastest-moving dials were presented in the middle in a head-locked manner. The performance here was considerably higher compared to condition B (48.2%), where large eye movements were necessary because the dials were located in the corners. Condition C, which allowed for head movement, also yielded a significantly lower performance (52.7%) than condition A. The other world-locked conditions (E to J) also showed a relatively low average performance, from 33.9% to 43.5%.

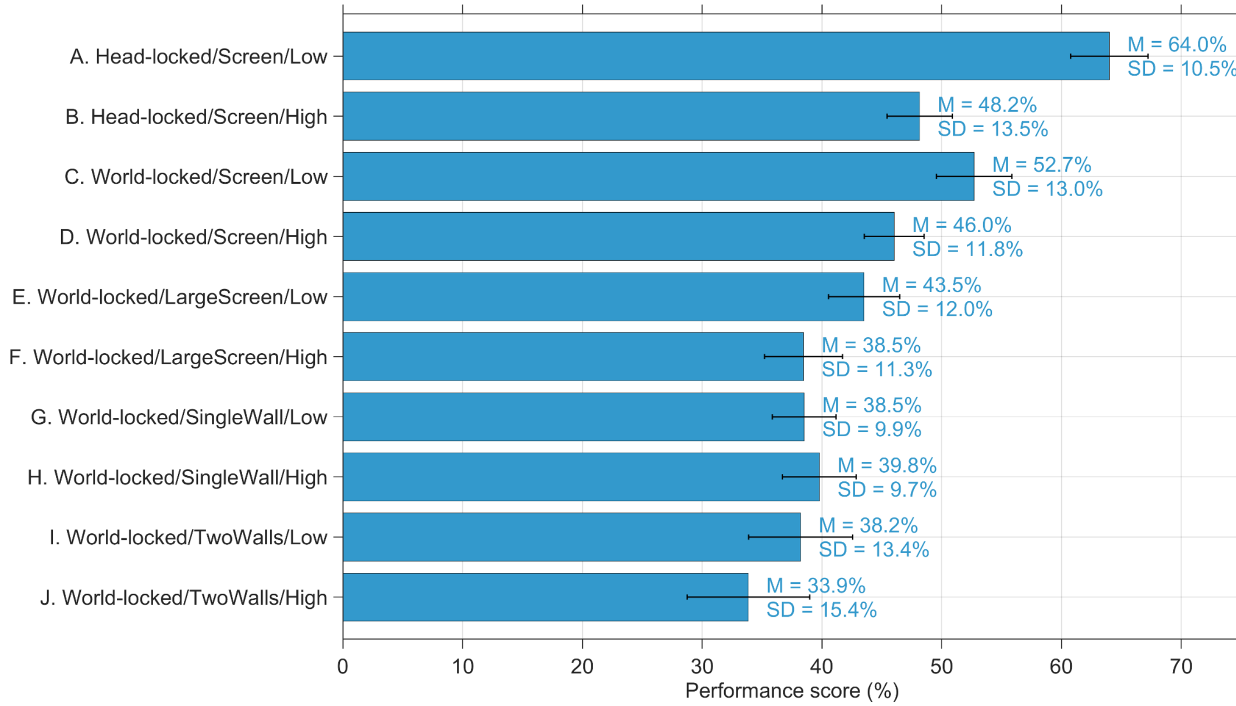


Figure 2. Means, standard deviations, and within-subject 95% confidence intervals for the performance score for the 10 conditions.

The percentage distribution of visual attention (Figure 3) showed that the faster-moving dials (0.32 Hz, 0.48 Hz) attracted the most attention, consistent with Senders' (1964) theory. This was particularly true when the faster-moving dials were located in the center on a central screen (conditions A, C, E). When these dials were positioned in the corners (B, D, F, H, J), attention distribution was more uniformly distributed across the six dials. The corresponding slopes of a linear regression line relative to bandwidth (Figure 4), where a higher slope can be seen as representing better attention distribution, are steepest for the low-effort configurations.

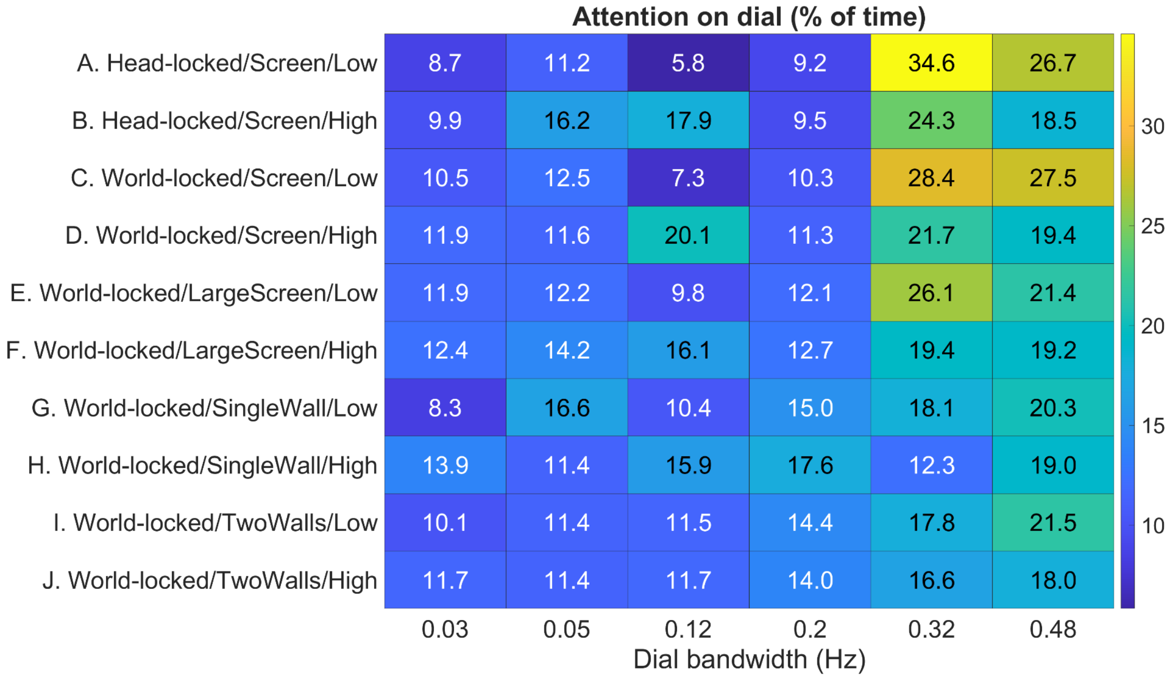


Figure 3. Means of 'Attention on dial' as function of dial bandwidth, for the 10 conditions.

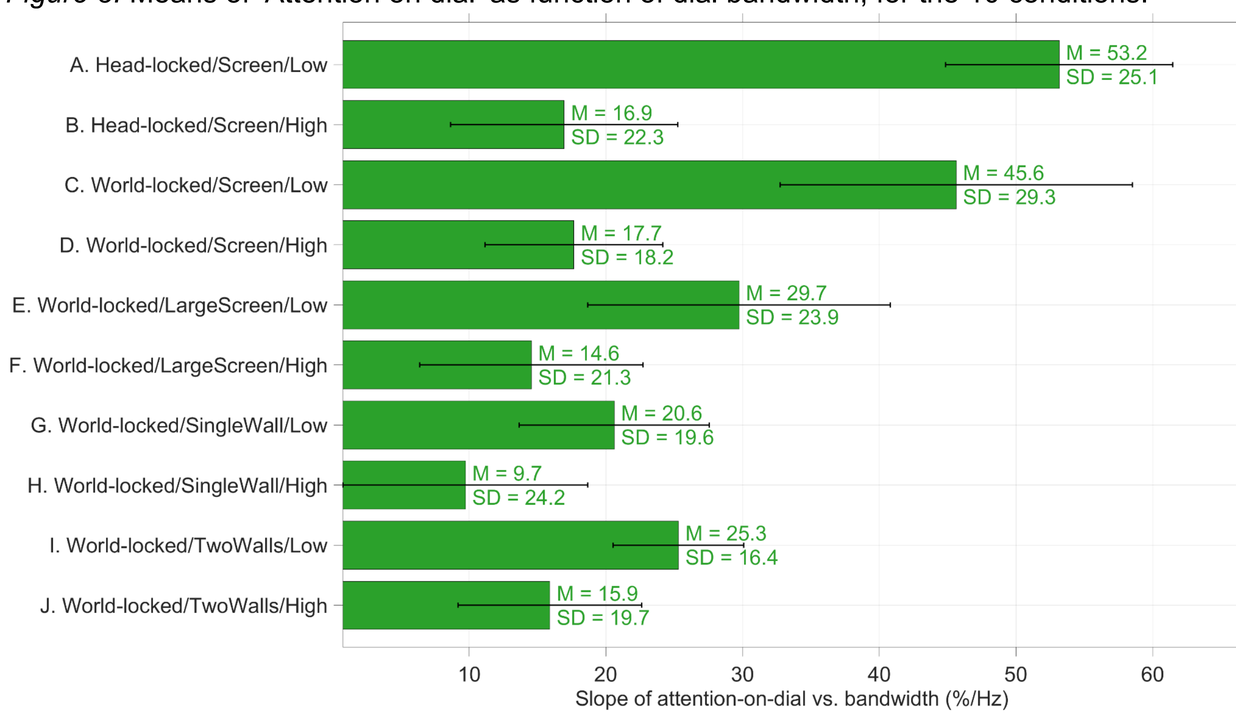


Figure 4. Means, standard deviations, and within-subject 95% confidence intervals for slope of participants' linear-fit regression lines of attention on dial versus dial bandwidth.

Figures 5, 6, and 7 show the amount of eye movement, head rotation, and head translation, respectively. It is noted that eye-movement effort is self-reported, whereas the head-movement variables are calculated from the orientation and position of the AR device, respectively. It can be seen that perceived eye-movement effort is similar across the dynamicity levels, but did differ between the low- and high-effort video configurations. That is, in the conditions where the dials were located in the corners, participants experienced higher effort, a replication of earlier findings in Eisma et al. (2018). The difference in eye-movement effort between the low- and high-effort variant was statistically significant, i.e., $p < 0.005$, for condition A versus B ($t(29) = -3.08$, $p = 0.004$), and for C versus D ($t(29) = -3.36$, $p = 0.002$), but not for E versus F ($t(29) = 0.16$, $p = 0.873$), G versus H ($t(29) = -0.38$, $p = 0.708$), and I versus J ($t(29) = -2.16$, $p = 0.039$).

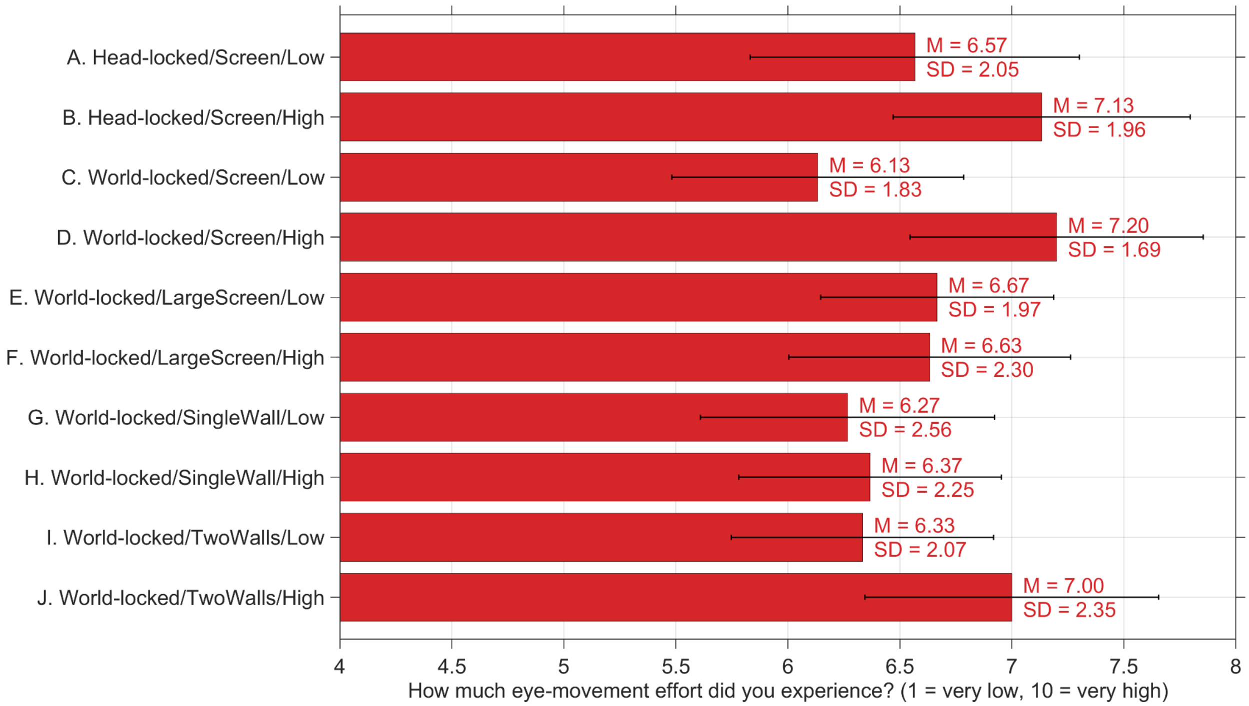


Figure 5. Means, standard deviations, and within-subject 95% confidence intervals for self-reported effort for the 10 conditions.

In terms of angular head movement activity (Figure 6), a pattern emerged where the amount of movement increased with dynamicity. The least head movement was observed in the head-locked conditions (A and B), which is expected as head movement served no purpose for the participants, i.e., it had no effect on what they saw. Significantly larger head-movement activity occurred in conditions E–H, where participants viewed a large screen. The highest head-movement effort was recorded in conditions I and J, where participants rotated their torsos and heads to read two screens on either side of the room. The total distance walked (Figure 7) followed a similar pattern. The option to stand, in particular, ensured that participants travelled over 6 meters on average.

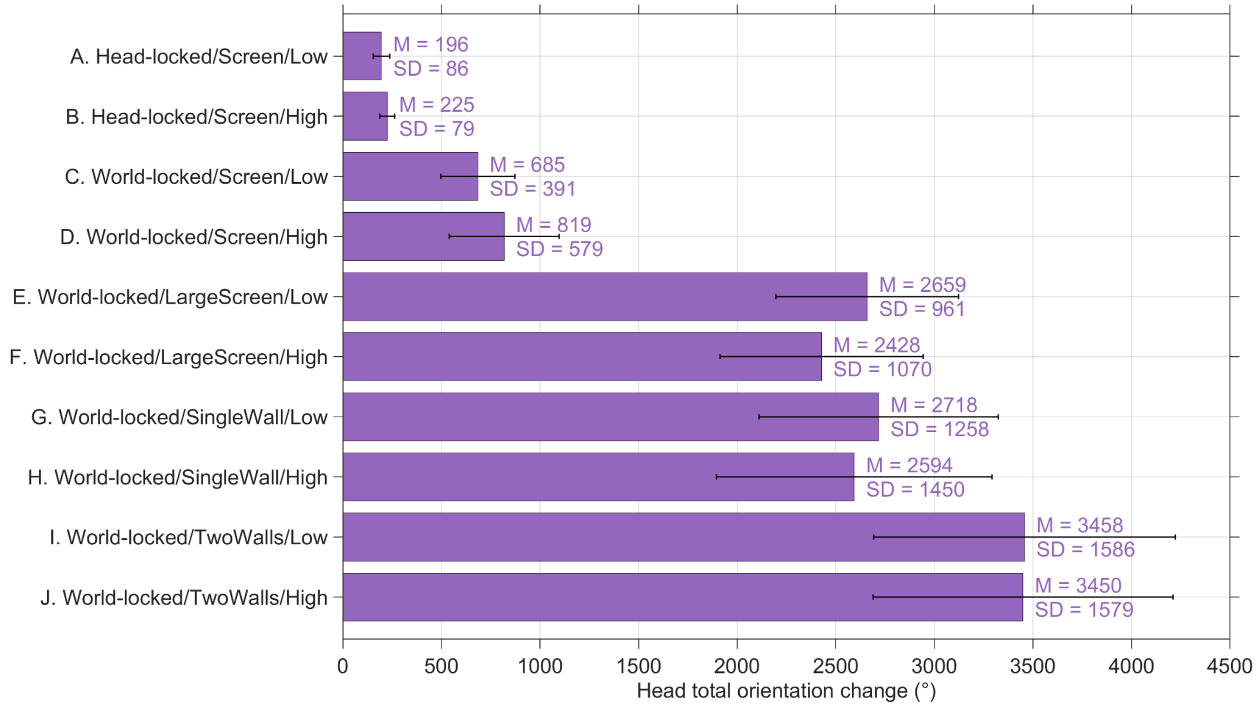


Figure 6. Means, standard deviations, and between-subjects 95% confidence intervals for total orientation change of the head for the 10 conditions. Result based on 19 of 30 participants.

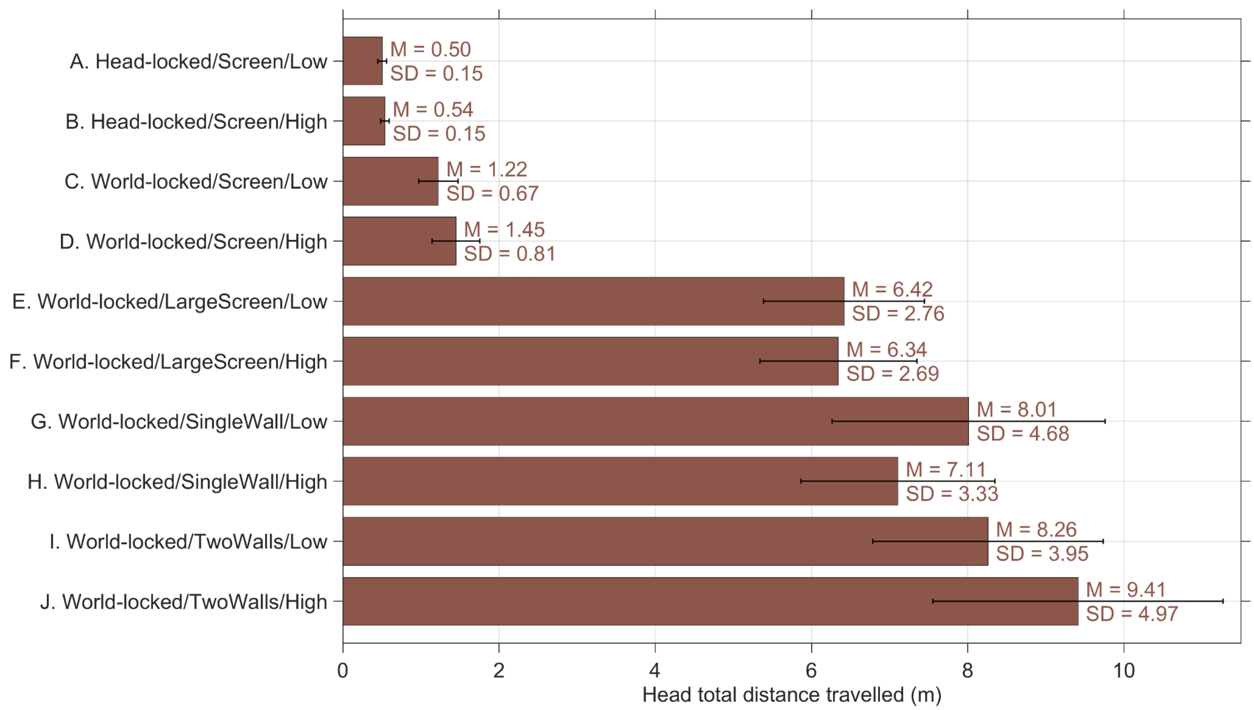


Figure 7. Means, standard deviations, and between-subjects 95% confidence intervals for total distance travelled of the head for the 10 conditions.

The transcripts (for Gemini-based summaries, see Appendix A) indicate that the ordinal progression from A/B to I/J expanded the monitoring behavior from purely eye-based scanning on a head-locked screen (A & B), to incorporating additional (small) head movements while seated before a world-locked display (C & D), and finally to relying on large required head, body, and ambulatory movements to scan increasingly distributed layouts on one (E–H) or two walls (I–J). This shift also saw participants move from being able to monitor all dials to adopting 'divide and conquer' strategies, where they would focus on a subset of dials on one side before deliberately switching their entire body and attention to the other.

In the low-effort dial configurations (conditions A, C, E, G, I), the centrally-located high-bandwidth dials allowed participants to adopt a central-focus strategy. Conversely, the high-effort conditions (B, D, F, H, J) invalidated this approach by placing active dials in the corners. This forced participants into a more effortful scanning pattern that required large movements between the widely separated targets; it also often led to a feeling of being overwhelmed or missing events.

Discussion

This research aimed to investigate the effect of display characteristics (including spatial arrangement and anchoring) and graduated levels of embodied interaction freedom (through changes in posture and allowed movement), made possible using AR, on human visual sampling behavior, threshold-detection performance, and bodily activity during multi-dial monitoring. We quantified the transition from primarily oculomotor sampling to the use of head and body rotating, and, to a limited extent, ambulation.

Performance in detecting threshold crossings was highest (64.0%) in Condition A, characterized by a head-locked screen with centrally-located high-bandwidth dials, and seated participants with restricted movement. This performance score is high compared to a previous study by Eisma et al. (2018), which used a chinrest and a similarly sized screen (31° horizontal angular range vs. 36° in the present study) and reported an average performance of 52.9%. In contrast, the high-effort configuration of the head-locked screen in the current study (Condition B) resulted in a performance score of just 48.2%, similar to 48.1% in the Eisma et al. (2018) study. A possible explanation for the high score in Condition A is that the head-locked presentation might have encouraged participants to maintain a central gaze. Apart from the fact that the fastest-moving dials were positioned in the center, thereby increasing the likelihood that their threshold crossings would be detected, there may also be a strategic advantage to keeping one's gaze centrally focused: peripheral vision can then be used to detect any threshold crossings of directly adjacent dials (Eisma et al., 2024).

Introducing head movement freedom with a seated world-locked screen (Conditions C & D) also led to decreased performance. This occurred even though these conditions presented a world-anchored screen, which offered the option of head movement. An explanation lies precisely in this increased freedom, i.e., the moving of the head could inadvertently cause distraction by enabling participants to look at slow-moving dials. Performance was similarly reduced for standing conditions with larger or more spatially distributed world-locked screens(Conditions E–J). This outcome is in line with our central hypothesis that increased embodied interaction freedom and information distribution (as systematically varied across our conditions) in monitoring tasks requires compensatory physical movements and would involve less efficient attention allocation.

Our findings are consistent with previous research indicating that while head-locked displays can be experienced as unpleasant, they can offer a performance advantage in certain conditions (Ghasemi et al., 2021; Peereboom et al., 2024). For example, Ghasemi et al. (2021) found that with head-locked content users do not have to spend extra effort looking between a fixed AR screen and their keyboard. On the other hand, Fukushima et al. (2020) found that world-anchored text is more readable than head-anchored text while walking. One likely explanation is that the vertical shock from walking creates a visual blur for head-locked text that the brain cannot suppress. In other words, while a head-locked presentation may be unnatural and uncomfortable, a benefit is that presenting vital information right in front of the user's central field of view can yield improved task performance.

Another main finding, consistent with Senders' theory, is that faster-moving (i.e., higher bandwidth) dials generally attracted more visual attention. However, the efficiency of this attention allocation, as measured by the slope of attention versus bandwidth, was highest when high-bandwidth dials were central and movement was restricted (condition A). This efficiency tended to decrease as freedom of movement increased, such as with the introduction of head movements in seated conditions (C & D), indicating a less strict adherence to bandwidth-driven sampling. This suggests that the increased physical effort required for navigation and reorientation in these conditions diverted attentional resources away from optimal attention prioritization based purely on dial bandwidth.

Additionally, as hypothesized, increasing freedom of movement led to a progressive increase in objective measures of head rotation and movement. Participants shifted from predominantly eye-based scanning in head-locked conditions to incorporating head and, where physically possible, full-body reorientations in conditions allowing standing and walking. Self-reported eye-movement effort, however, was more strongly influenced by the dial configuration (i.e., high-effort corner placements increasing perceived effort than by the level of interaction freedom when viewing a relatively small screen. This suggests that engaging the 'head field' alleviates the physical strain on the eyes by reducing the need for extreme eye movements to access eccentric objects. Qualitative reports corroborated the objective measures, detailing a shift from purely eye-based scanning to incorporating head turns, torso rotation, and ambulation, often adopting a 'divide and conquer' approach for the distributed screens.

Based on the research findings, it can be recommended to minimize the physical effort that an operator must exert. That is, critical information should not be distributed across large screens or throughout a physical space, as the results demonstrate that such setups do not lead to better performance, even when operators are granted the freedom of head movement. Instead, the most important information should be presented centrally and saliently, as is done with a Safety Parameter Display System (SPDS), process safety dashboards, or command and control displays in process control or command centers (e.g., Sheridan, 2021). In our case, the best performance was achieved in a setup where the most vital information was forced in the center of the field of view.

Several limitations should be acknowledged. First, movement in the ambulatory conditions (E–J) was limited by the 3-m data cable. Second, while conditions E–J were designed to simulate large screens, conceptually equivalent to a control room meta-display, the experiment was conducted in a relatively small room. This mismatch likely reduced participants' perception of the sizes of the screens. Because the screens in conditions G–J were large and positioned far away, walking towards the screens had little functional impact on improving the participants' view. The measured ambulatory movement should therefore be interpreted as an indicator of participants' *inclination* to walk, rather than a strategy to gain a better vantage point. Third, while the Magic Leap 2 is a see-through AR headset, its digital field of view (45°H × 54°V) is narrower than natural human vision. This meant that for wider screen configurations, digital content extending beyond this area required head movements to be brought into view, which could have influenced scanning behaviors. Fourth, the dial-monitoring task, while a well-established paradigm, is abstract and may not fully generalize to all complex operational environments. For example, this study focused exclusively on monitoring virtual information and did not include tasks requiring participants to interact with or monitor the physical world visible through the headset. Finally, the participant sample, while adequate for this study (*n* = 30), consisted primarily of university-affiliated individuals who were largely familiar with AR/VR technology, which may limit the generalizability of the findings to other populations.

Conclusion

This study, using augmented reality to examine information distribution and embodied interaction, offers insights into how people sample visual information. We found that reducing physical constraints on embodied interaction and centralizing spatial information distribution, particularly by placing high-bandwidth information centrally, yielded the best detection performance and most effective attention distribution. Human operators naturally move to gather information from spread-out displays, but these movements appear to be a compensatory strategy to meet the increased physical requirements of accessing dispersed information, rather than a path to more efficient attention. For designers of AR and control room interfaces, this means key information should be central to minimize physical effort and optimize mental resources

References

AssemblyAI. (2025). API reference. AssemblyAI. https://www.assemblyai.com/docs/api-reference

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6–10. https://doi.org/10.1038/s41562-017-0189-z

Cumming, G. (2009). Inference by eye: Reading the overlap of independent confidence intervals. *Statistics in Medicine, 28*, 205–220. https://doi.org/10.1002/sim.3471

Eisma, Y. B., Bakay, A., & De Winter, J. (2024). Expectancy or walience?—Replicating Senders' dial-monitoring experiments with a gaze-contingent window. *Human Factors, 66*, 1770–1785. https://doi.org/10.1177/00187208231176148

Eisma, Y. B., Cabrall, C. D. D., & De Winter, J. C. F. (2018). Visual sampling processes revisited: Replicating and extending Senders (1983) using modern eye- tracking equipment. *IEEE Transactions on Human- Machine Systems, 48*, 526–540. https://doi.org/10.1109/THMS.2018.2806200

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*, 24–29.

Fukushima, S., Hamada, T., & Hautasaari, A. (2020). Comparing world and screen coordinate systems in optical see-through head-mounted displays for text readability while walking. *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Porto de Galinhas, Brazil, 649–658. https://doi.org/10.1109/ISMAR50242.2020.00093

Gasse, A., Lingler, A., Lorenz, M., Oulasvirta, A., Wintersberger, P., & Ebel, P. (2025). Evaluating attention management systems for dynamic monitoring tasks. *CHIWORK '25 Adjunct: Adjunct*

Proceedings of the 4th Annual Symposium on Human–Computer Interaction for Work (pp. 1–6). Association for Computing Machinery. https://doi.org/10.1145/3707640.3731920

Ghasemi, Y., Singh, A., Kim, M., Johnson, A., & Jeong, H. (2021). Effects of head-locked augmented reality on user's performance and perceived workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*, 1094–1098. https://doi.org/10.1177/1071181321651169

Hancock, P. A., Hancock, G., Sellen, A., Lee, J., Sawyer, B., Sheridan, T., Milgram, P., & Sanderson, P. (2019). The life and legacy of Professor John Senders. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*, 577–581. https://doi.org/10.1177/1071181319631191

Kovesdi, C., Spielman, Z., LeBlanc, K., & Rice, B. (2018). Application of eye tracking for measurement and evaluation in human factors studies in control room modernization. *Nuclear Technology*, *202*, 220–229. https://doi.org/10.1080/00295450.2018.1455461

Le Blanc, K., Joe, J., Rice, B., Ulrich, T., & Boring, R. (2015). *Benefits of advanced control room technologies: Phase one upgrades to the HSSL and performance measures* (INL/EXT-15-35381, Revision 1). Idaho National Laboratory. https://doi.org/10.2172/1196559

Magic Leap. (2022). Magic Leap 2’s advanced AR platform and revolutionary optics. https://www.tdsynnex.com/na/us/magic-leap/wp-content/uploads/sites/85/2023/10/Whitepaper-Magic-Leap-2-Optics.pdf

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*, 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Peereboom, J., Tabone, W., Dodou, D., & De Winter, J. C. F. (2024). Head-locked, world-locked, or conformal diminished-reality? An examination of different AR solutions for pedestrian safety in occluded scenarios. *Virtual Reality, 28*, Article 119. https://doi.org/10.1007/s10055-024-01017-9.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*, 643–675. https://doi.org/10.1037/0033-295X.106.4.643

Poole, C. A., Warden, A. C., Wickens, C. D., Raikwar, A., Clegg, B. A., Buckman, M., & Ortega, F. R. (2025). *Information access costs with an augmented reality head-mounted display.* Manuscript in preparation.

Rouse, W. B. (1981). Human-computer interaction in the control of dynamic systems. *ACM Computing Surveys (CSUR)*, *13*, 71–99. https://doi.org/10.1145/356835.356839

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Transactions on Human Factors in Electronics, HFE-5*, 2–5. https://doi.org/10.1109/THFE.1964.231647

Senders, J. W. (1983). *Visual sampling processes* (Doctoral dissertation, Katholieke Hogeschool Tilburg, The Netherlands).

Sheridan, T. B. (2021). Human supervisory control of automation. In G. Salvendy & W. Karwowski (Eds.), *Handbook of human factors and ergonomics* (5th ed.). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119636113.ch28

Sidenmark, L., & Gellersen, H. (2019). Eye, head and torso coordination during gaze shifts in virtual reality. *ACM Transactions on Computer-Human Interaction, 27*, 1–40. https://doi.org/10.1145/3361218

Tabone, W., & De Winter, J. C. F. (2023). Using ChatGPT for human-computer interaction research: A primer. *Royal Society Open Science, 10*, Article 231053. https://doi.org/10.1098/rsos.231053

Warden, A. C., Wickens, C. D., Clegg, B. A., Rehberg, D., & Ortega, F. R. (2024). Information access effort: The role of head movements for information presented at increasing eccentricity on flat panel and head-mounted displays. *Human Factors, 66*. https://doi.org/10.1177/00187208231204567

Wickens, C. D. (2008). *Visual attention control, scanning, and information sampling.* In C. D. Wickens & J. S. McCarley (Eds.), *Applied attention theory* (pp. 41–61). CRC Press. http://doi.org/10.1201/9780429059261-4

Appendix A

The prompt used was as follows:

Summarize the interview transcripts for the attached 10 conditions. Conditions B, D, F, H, and J are the high-effort configurations, with the fast-moving dials in the corners. In contrast, conditions A, C, E, G, and I are their equivalents but with the fast-moving dials located more centrally. Please summarize as follows, using two sentences for each pair:
A&B:
C&D:
E&F:
G&H:
I&J:

Table A1
Gemini 2.5 output based on participants' interview transcripts about the strategies they employed.

A&B: *In condition A, participants often kept their head relatively still, focusing their gaze on the center of the screen to monitor the faster-moving central dials while using peripheral vision for the others. For condition B, where the faster dials were in the corners, participants reported that keeping the head still was essential because the interface was head-locked, forcing them to use more demanding eye movements to scan between the unpredictable, far-apart dials.*
C&D: *For condition C, participants generally focused on the two fastest-moving dials in the middle, often leaning back to get a wider view and using a combination of head and eye movements. In condition D, participants also identified and focused on the fastest dials, but since these were now in the corners, it required more frequent and extensive head and eye movements to scan from side to side or from top to bottom.*
E&F: *In condition E, the large and spaced-apart dials necessitated significant head and body movement, with many participants stepping back to get a better overview and then focusing on the faster dials, which were generally in the center. Condition F amplified this challenge; due to the fast-moving dials being in the corners of an already large display, participants had to engage in even more extensive head, body, and eye movements, often walking or turning to cover the area between the active corners.*
G&H: *For condition G, where dials were on a side wall, participants would turn their body to face them and then primarily use head movements to scan the row, often focusing on the faster dials in the middle of the set. In condition H, this task became more demanding as the fastest dials were at the far ends of the row; this forced participants to use larger, more frequent head and body movements to scan the entire width of the display.*
I&J: *With dials on opposite walls in condition I, participants typically stood in the middle and switched their focus from one side to the other, primarily using head and upper-body rotation. Condition J was perceived as the most difficult, as the fastest dials were on opposite sides, requiring constant and physically demanding switching between the two walls with significant head and body movement, making it hard to monitor all dials effectively*