Multitasking Across Physical and Virtual Displays: The Effect of Spatial Discontinuity and Task Load

Peiyu Zhang, Mohamad El Iskandarani, and Sara Riggs

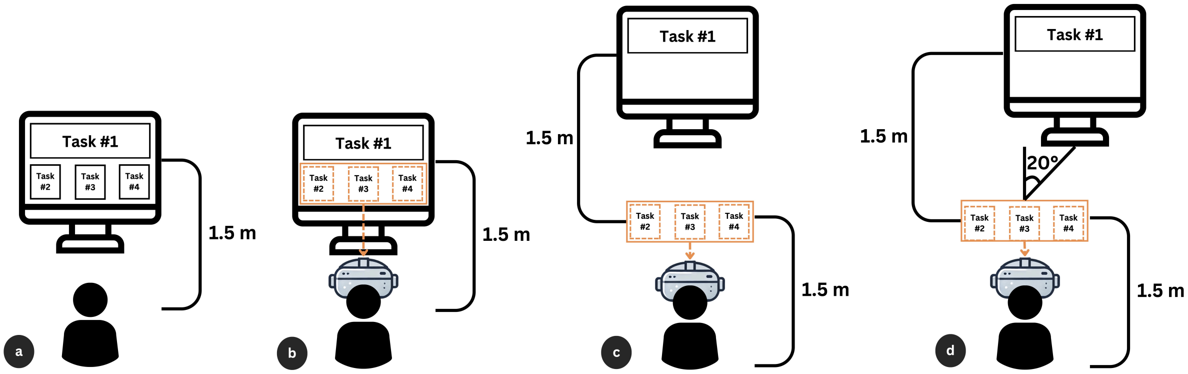

Fig. 1: Illustration of the four experimental physical/virtual configurations. Across all images, there are four tasks: Task #1 is the shooting task, Task #2 is system monitoring, Task #3 is tank monitoring, and Task #4 is parameter entry. The configurations evaluated included the following: (a) *"Desktop"* configuration where all tasks were displayed on a single desktop screen positioned 1.5 m away from the user; (b) *"Overlay"* configuration where the shooting task was shown on the desktop display and the other three tasks were on an augmented reality display positioned in the same location as the desktop; (c) *"Distant"* configuration where the desktop and AR displays were separated by 1.5 m to introduce depth discontinuity; and (d) *"Misaligned"* configuration where the desktop display was shifted 20 degrees to the right of the user's viewpoint, creating both depth and visual field discontinuities between displays. The visualizations are not drawn to scale and are intended for illustrative purposes only.

Abstract— Optical see-through head-mounted displays (OST-HMDs) allow users to interact with both physical and virtual displays overlaid in the real world within augmented reality (AR) environments. When arranging these displays in space, spatial discontinuity—such as differences in viewing distance or lateral separation in the visual field—may occur. This study examined the effects of displays' spatial discontinuities on multitasking performance and perceived workload in AR, as well as how these effects were influenced by task load. Participants completed a multitasking simulation with four tasks distributed across physical and virtual displays under four display configurations: two configurations with spatial discontinuity, one cross-display configuration without discontinuity, and a baseline configuration where all tasks were shown on the physical display. Each configuration was tested under two task load levels. Results showed that OST-HMD-supported AR significantly impaired multitasking performance, regardless of whether spatial discontinuity between virtual and physical displays was present. Introducing display discontinuities through small depth separations or alignment offsets produced no additional significant influence on performance. The AR configurations also shown increased perceived workload than the baseline, but spatial discontinuity did not significantly affect workload ratings. Task load modulated the effect of display configurations on performance as well as physical and frustration ratings of the perceived workload.

Index Terms—Augmented reality, multitasking, context switching, display discontinuity, multi-display user interface

1 INTRODUCTION

With advances in augmented reality technology, virtual content can now be overlaid on the real-world environment, often through OST-HMDs [10]. This capability supports a wide range of applications where users can simultaneously interact with both physical objects and virtual content in domains such as data analytics [18], maintenance [40, 3], and surgery [29]. Designers can relocate content traditionally shown on physical displays to AR displays to enhance user experience, as seen in applications like heads-up displays in cars [55] and the presentation of virtual photos and text for tourists [37]. When space is limited, replacing physical monitors with virtual displays shown on AR headsets can also offer greater flexibility [43]. Although AR offers many advantages, prior work has raised concerns about potential negative impacts on performance [43, 2, 19], comfort [28, 52, 27], and perceived workload [26, 21] caused by OST-HMDs. In multitasking scenarios, users often need to continuously monitor events across multiple screens. Because virtual screens in AR can be freely placed in space, spatial discontinuities may occur because of occlusion by physical obstacles, the presence of other displays, or design choices. For example, displays positioned at varying distances from the user can create depth discontinuity, and gaps between displays in the visual field can create visual field discontinuity [47]. These discontinuities could potentially impose additional burden on switching between displays.

In this paper, we aim to provide design insights for multitasking contexts involving a hybrid setup of physical and OST-HMD rendered virtual displays, focusing on spatial discontinuity between displays under time constraints. To this end, we compared task performance and perceived workload in an AR-based multitasking simulation with

---

Peiyu Zhang, Mohamad El Iskandarani, and Sara Riggs are with the Department of Systems and Information Engineering, University of Virginia, Charlottesville, USA.
E-mail: {pz3sf, anz8av, sriggs}@virginia.edu

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

a baseline in which no AR was used. Within the AR environment, there were three display configurations: “overlay” with no spatial discontinuity, “distant” with depth discontinuity only, and “misaligned” with combined depth and lateral visual field discontinuity, as shown in fig. 1. Because multitasking contexts such as control rooms and maintenance work often involve fluctuating levels of task demand, we also examined whether these effects could amplify under higher task load by manipulating event frequency and time pressure. Our findings show that the hybrid setup of physical and virtual displays led to a decline in performance and an increase in perceived workload compared to multitasking with only a physical display. Introducing a small depth separation and lateral misalignment between displays did not significantly impact performance or workload overall. The differences in workload and the number of errors made between configurations were more pronounced under higher task load, whereas the differences in response time were less pronounced. We also offered design guidelines derived from the insights of our results.

2 RELATED WORK

2.1 Real-World Applications of AR with OST-HMDs

OST-HMD–supported AR has been applied across a wide range of domains. In medicine, AR projects anatomical or diagnostic imagery directly onto patients or surgical sites to guide planning and procedures [17]. In architecture and construction, OST-HMDs visualize digital building models in context to support design and facility management [59]. In industrial settings, AR overlays assembly instructions and technical diagrams onto machinery to aid diagnostics, maintenance, and production [35]. Military training incorporates AR to present tactical information and terrain visualizations for exercises [8], while archaeology uses AR headsets in the field to reveal excavation layers and support documentation [14]. Beyond these domains, many applications require users to divide attention across physical and virtual displays. For office and mobile work, AR virtual monitors extend traditional screen space by adding floating displays alongside laptops or desktops [42, 41, 11]. In collaborative environments, large wall displays are paired with OST-HMDs so that shared content appears physically while personal tools remain virtual [48]. AR overlays also aid navigation during walking by providing real-time directional cues [39]. Collectively, these examples show how OST-HMD–supported AR complements physical screens, enabling hybrid display setups that adapt flexibly across work, collaboration, and everyday activities. In many of these examples, such as surgery and military training, task load may shift across scenarios, which should be considered in design.

2.2 Challenges in OST-HMD–based AR

To gain a clear view of an object, our eyes converge based on depth cues (vergence) and the lens adjusts focus accordingly to the focal plane (accommodation). When looking at a physical display, the depth cues align with the focal plane determined by the actual distance of the display, thus aligning the vergence and accommodation. However, when looking at a virtual display on common OST AR headsets (like the one in this study), depth cues are not aligned with the actual focal plane in most cases [30]. Although the headset can simulate different depth cues, it has a fixed focal plane. The misalignment of depth cue and focal plane leads to a conflict between vergence and accommodation. Previous research has demonstrated that this misalignment distorts depth and size perception [31] and can result in performance decrement and visual fatigue [23].

More importantly, even if the physical and virtual displays were placed at the same depth in AR, our eyes still need to adjust accommodation when switching gaze between them. This is because physical displays exist at their true optical depth, while virtual displays remain anchored to the fixed focal plane of the headset. Prior studies have demonstrated that such *focal plane switching* can impair performance and increase eye fatigue in a visual search task, with both monocular and binocular AR displays, especially when the transition between displays leads to blurring of the content shown [19, 2]. Chiossi et al. [13] used event-related potential (ERP) markers to show that visual discomfort caused by blur could impose high neural costs. Furthermore, even when focal distance and the simulated depth of virtual displays are matched to those of physical displays—–thus eliminating the need for focal plane switching—–studies have found that *context switching* between the physical and virtual environments can still increase eye fatigue with both binocular and monocular OST displays [19, 2]. However, context switching generally did not impact performance on virtual search tasks [19, 2, 16], except when the virtual content became too distant and began to appear blurry [19].

Beyond these perceptual and visual challenges, there are also physical and ergonomic issues to consider with AR headsets. Unlike desktop monitors, they add weight and pressure to the head and neck, leading to greater strain and fatigue [28]. AR use also increases muscle activity and ergonomic risks [36], with extended sessions linked to muscular fatigue and reduced awareness in workplaces [51]. Neck muscle contraction further adds cumulative strain [60]. Stereoscopic content outside the ‘zone of comfort’ causes eye fatigue [50], and AR setups often induce nausea [52], motion sickness [27], and dizziness [51].

2.3 Spatial Arrangement of AR Displays

Prior work has examined how the spatial arrangement of displays in AR affects user interaction. In collaborative sensemaking, surfaces and items in the physical environment guided how virtual documents were arranged [34]. In hybrid setups with mobile devices and OST-HMDs, layout strategies that varied text size and spacing influenced readability, comfort, and speed, reflecting trade-offs between viewing more content and maintaining legibility [4]. Body-relative placements also strongly affect interaction: interfaces near eye level or on-body were most convenient, while world-anchored views better supported spatial understanding [24, 37]. Head-locked overlays increased throughput, whereas world-locked overlays were rated more usable [20]. During walking, head anchoring preserved locomotion and task performance, while hand anchoring slowed walking but reduced missed responses [46]. On-hand reference frames showed that pinch-offset placements improved speed and accuracy but introduced occlusion and fatigue [25]. Direction and eccentricity also matter: task cues were most effective in the lower visual field at small eccentricities, while non-task information worked better in the left or top-left [57]. For alerts, object-centered placements improved speed and lowered workload [45], while peripheral placements reduced disruption but risked misses [44]. Adaptive strategies aligning content with environmental context enhanced efficiency and usability, whereas stable anchoring improved precision in dynamic tasks [15, 1].

Like other systems that involve multiple physical displays, hybrid configurations of virtual and physical displays are multi-display user interfaces (MDUIs). In the arrangement of such MDUIs, each pair of displays can exhibit *spatial discontinuity*, which may arise from occlusion by obstacles in the physical environment, the presence of other displays that block portions of space, or additional design considerations. Such cases were also evident in the work discussed in the previous paragraph. Rashid et al. [47] distinguished two primary forms of spatial discontinuity (1) *depth discontinuity*, where displays are positioned at different distances from the user, and (2) *visual field discontinuity*, where displays are separated in the visual field, regardless of whether they are on the same depth plane. Prior research has shown that both types of discontinuity can degrade performance [58, 38, 7], although the extent of their effects varies depending on the task and display type [47]. The angle of visual separation also matters, as Warden et al. [53] found that increasing the separation up to 40 degrees between two numbers did not affect response time or accuracy in either a digit reading task or a mental subtraction task, but larger angles above about 32 degrees required more head movement, which could potentially have a greater impact on more demanding AR tasks.

According to Multiple Resource Theory [54], task cues such as salient color or motion can be detected through ambient vision, drawing on a different pool of visual resources than tasks that require focal vision, such as reading. However, when the ambient-vision task is placed farther from the user, the cue may no longer be detectable peripherally and instead require focal vision, creating competition with focal tasks and potentially degrading performance. In the context of switching between physical and virtual displays in AR, increasing the depth separation between physical and virtual displays may further widen the gap between their respective focal planes. The resulting longer distance of focal plane switching has been shown to continue elevating eye fatigue and impair performance in visual search tasks [2]. However, little research has examined the combined effect of depth and visual field discontinuity in AR settings requiring continuous monitoring of both physical and virtual displays.

3 PRESENT WORK

Building on prior work, we identified a research need: examining the effect of spatial discontinuity on multitasking in environments that require continuous monitoring of both virtual and physical displays. We also varied task load, as Perceptual Load Theory [32] suggests that increasing perceptual demands reduces the capacity available for additional tasks, and could make concurrent monitoring more difficult under high load.

We designed a set of tasks with frequent events requiring rapid responses across displays, which we hypothesized would amplify the impact of factors such as context and focal plane switching compared to prior work. Fig. 2b shows the normal (no-event) appearance of all tasks to illustrate the default UI layout, and different event states are illustrated in Figures 3-5. Compared to earlier studies evaluating performance on hybrid displays in AR, our study required more frequent switching between displays by incorporating multiple monitoring tasks with high response frequency and assigning different tasks to each display. The simulation included four configurations shown in fig. 1. In the *desktop* configuration, all tasks were on a physical monitor. The *overlay* configuration moved tasks to a virtual display on an OST-HMD without spatial discontinuity, isolating burdens introduced by AR, such as context and focal plane switching. Notably, the focal plane distance difference in this configuration (0.76 m) was smaller than in previous studies [2, 19]. The *distant* configuration introduced depth discontinuity, while the *misaligned* configuration further added lateral misalignment to create visual field discontinuity. Finally, we examined how task load—manipulated through event frequency and time pressure— modulated the effects, and collected the NASA-TLX [22] responses to capture perceived workload. Together, these configurations aimed to provide insights into whether the spatial flexibility offered by AR justifies potential trade-offs in user performance and workload.

Our research questions were as follows:

1. How does spatial discontinuity between displays affect task performance and perceived workload in multitasking AR?
2. How does task load (low vs. high) modulate these effects?

And our expectations were:

1. Compared to the desktop baseline, the AR configurations would decrease performance and increase workload due to the burdens introduced by the AR use.
2. The distant and misaligned configurations, which introduced spatial discontinuity between displays, would further decrease performance and increase perceived workload compared to the overlay and baseline physical-only configurations.
3. Higher task load would make differences in performance and workload across configurations more pronounced.

4 METHODS

4.1 Apparatus

Fig. 2a shows a picture of the experiment setup in the overlay configuration. The virtual display was shown on a Magic Leap 2 see-through AR headset with a resolution of 1440 × 1760 pixels per eye and a diagonal FOV of 70 degrees. The focal plane of the headset is at 0.74 m. The desktop display used to present the search target was a 27-inch monitor with a resolution of 2560 × 1440 pixels, connected to a computer. The experiment used a simulator with four tasks developed with the Unity game engine. Participants interacted with the tasks using a three-key mini Bluetooth keyboard (held in their left hand) and a Bluetooth mouse (operated with their right hand), both placed on a fixed-position table. They remained seated in a stationary chair throughout the session. The physical desktop display was on another table, which was repositioned accordingly depending on the experimental configurations described in fig. 1.

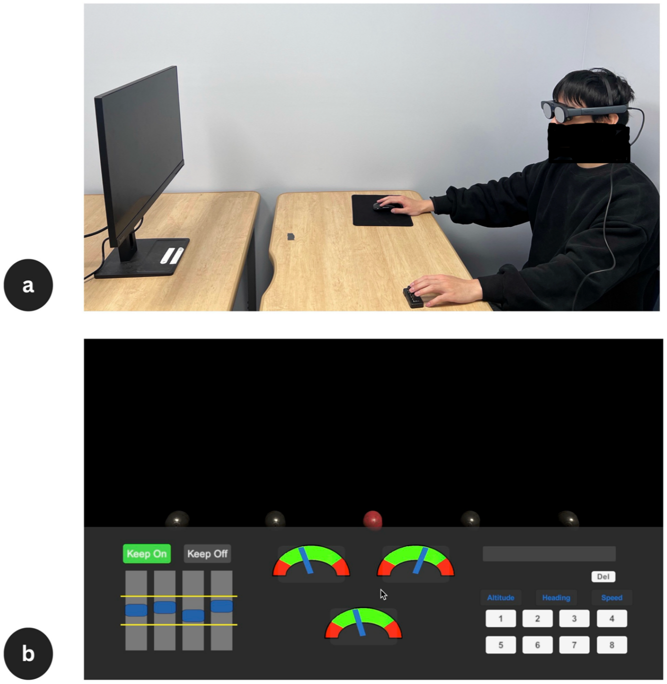
[redacted]

Fig. 2: (a) The experimental setup in the overlay configuration with the participant wearing the AR headset and using the three-key mini keyboard with their left hand and mouse with their right hand. (b) A screenshot of the simulation environment with the four tasks at normal (no-event) states. Different event states for each task are in fig. 4, fig. 3, and fig. 5.

4.2 Participants

Twenty university students (10 males, 10 females, M=21.3, SD=2.7) were recruited for this study. All participants were right-handed and had normal or corrected-to-normal vision and normal color vision. The study was approved by the University of Virginia's Institutional Review Board. Standard IRB protocols were followed to ensure informed consent, participant anonymity, and data security throughout the experiment. The participants were compensated with gift cards for their time.

4.3 Tasks

The tasks in the simulator were inspired by the NASA Multi-Attribute Task Battery-II [49], a widely used tool for assessing multitasking performance. We modeled a control-room–style situation in which an operator must attend to multiple concurrent subtasks and occasionally interact with each. The tasks we evaluate are independent task streams (each generates its own events and performance measures) and were designed to span different and distinct task types: continuous monitoring and responding with salient visual cues (i.e., system monitoring and tank monitoring), a discrete multi-step update task requiring focal attention and short-term retention (i.e., parameter entry), and a time-pressured response stream that induces interruption pressure and frequent cross-display switching when presented on a separate display

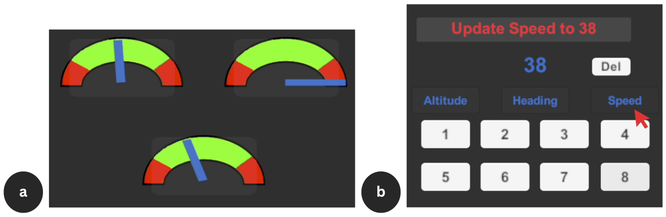

Fig. 3: (a) A screenshot of an event state of the tank monitoring task where the dial in the top right gauge entered the red zone. (b) A screenshot of an event state of the parameter entry task with a prompt to, “Update Speed to 38”. Participants would need to click on 3, 8, and then the Speed button to fulfill this request.

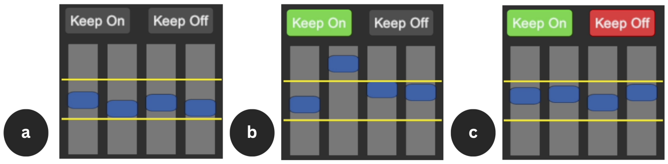

Fig. 4: Screenshots of the system monitoring task: (a) An event state with the “Keep On” button turned off (i.e., not green). (b) An event state that had a floating bar above the yellow line (bar second to the left). (c) An event state with the “Keep On” button turned on (i.e., red).

(i.e., shooting). Fig. 2b provides an overview of the simulation environment and experiment setup, and the details of each sub-task are described below.

The simulation was dynamic in nature, and each task had a normal state, as shown in fig. 2b, and at least one kind of event state requiring responding with a mouse or keyboard. There was no designated primary task, and users were instructed to respond to all tasks as quickly and accurately as possible. A pilot study ($n = 5$) was conducted to validate the task design and calibrate event frequency for low and high task loads, confirming that after tasks were moved from the physical display to AR, their size and mouse speed remained consistent with the physical setup, and that the physical screen stayed clearly visible across repositioned display configurations.

The **tank monitoring** task required the user to monitor three dials, which updated to a random position every 0.5 seconds. The user needed to make sure all three dials stay within the green range. If the needle of a dial entered one of the out-of-range red zones, such as shown in fig. 3a the user must click the gauge to reset it. Only one dial became abnormal at a time. After making the correct selection, the task would reset to the normal state.

The **system monitoring** task was similar to tank monitoring in nature, but consisted more components that the user must monitor: A “Keep On” button, a “Keep Off” button, and four floating bars. If the “Keep On” button changed color from green to gray, as shown in fig. 4a, or the “Keep Off” button changed color from gray to red, as shown in fig. 4c, the user were tasked to click on it. In short, the user should ideally keep the ‘Keep On’ button green and the ‘Keep Off’ button gray. Additionally, if any of the four floating bars moved outside the region between the two yellow lines, as shown in fig. 4b, the user must click anywhere in the floating bar area to restore it to the correct position. After making the correct click, the task would return to the normal state.

The **parameter entry** task required the user to follow a command displayed on the screen, instructing them to update one of three parameters—attitude, heading, or speed, as shown in fig. 3b. To enter the requested value, the user clicked the corresponding numbers on a number pad and then selected the appropriate parameter button. If an incorrect number is entered, the user can press the “Del” button to backspace. The number would always have two digits. After clicking the parameter button, the task would return to the normal state regardless of the correctness.

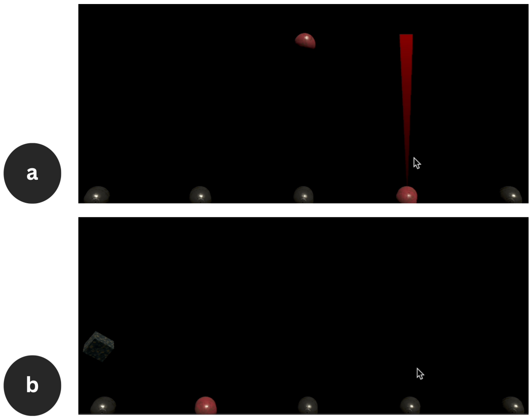

Fig. 5: Screenshots of the shooting task: (a) An event state where the target (i.e., red sphere target) appeared and needed to be eliminated. (b) An event state where there was a false target (i.e. black cube) appeared and did not require any action from the user.

The **shooting** task shown on fig. 5 required the user to shoot at incoming targets that descended downward from the top of the display. Unlike the other three tasks, the user responded to events from this task using a mini keyboard that had three keys. The user could toggle between the five different cannons along the bottom of the task by pressing the left or right key on the keyboard, and the middle key was used to shoot a laser from the selected cannon. The active, selected cannon would be highlighted red while the others remain black. When a target descended from the top and it was a red sphere, as in shown fig. 5a, the user must align the correct cannon and shoot before the target reached the bottom. However, if the target was not a red sphere such as the case shown in fig. 5b, then the user did not need to take any action. When the target was eliminated, regardless of correctness, or when the target reached the bottom, the event would end and the task would return to the normal state.

4.4 Experimental Design

There were four display configurations, as shown in fig. 1. Participants completed tasks under low and high task load for each configuration—resulting in a total of eight conditions per participant. Participants were trained with a task load set between the high and low conditions before the experiment. Training included both the desktop and overlay configurations, lasting at least three minutes per configuration and extended as needed until participants felt confident. After completing each condition, the participants filled a NASA-TLX questionnaire to assess their workload. The system automatically documented the response time for each event and the errors made.

The sequence of conditions was randomized for each participant. In the **desktop** configuration, all tasks were displayed on the desktop monitor, placed 1.5 m from the user. In the **overlay** configuration, the system monitoring, tank monitoring, and parameter entry tasks were moved to the virtual display but remained in the same depths and sizes as in the desktop configuration. Because the AR headset had a hardware-fixed focal plane at 0.74 m, this created a 0.76 m difference between the focal planes of physical and virtual displays. In the **distant** configuration, the desktop display was moved 1.5 m away from the user, while the virtual display remained in the same location. In the **misaligned** configuration, the desktop display was shifted 20 degrees to the right of the virtual display, which remained fixed in place. The misalignment was intentionally kept small enough that users could look back and forth between displays without requiring head movement, as validated in the pilot test, ensuring that tasks did not move outside the field of view and introduce an additional factor influencing responses. And while head movement was not prohibited, participants were encouraged to minimize it. The virtual display had a dark gray background identical to that of the physical display to ensure tasks were clearly visible in all configurations, and it was world-fixed with a constant size across configurations. Participants used the same input devices (mouse and keyboard) in the all configurations.

Each condition lasted three minutes and followed a pre-generated input file to determine task events. Eight different input files were used across conditions to prevent participants from memorizing the events, but the same set of input files was used across all participants. The number of events and the response time window for each task varied by task load. Events from system monitoring, tank monitoring, and parameter entry never occurred overlapped, but shooting task events could coincide with one of these tasks to induce switching between displays. Therefore, the shooting task required more frequent responses than the other three tasks and was on a separate display, and among those three, the parameter entry task took more time to response as it required three clicks.

Within an input file, there would be an equal number of events for the system monitoring, tank monitoring, and parameter entry tasks, with the order of events randomized. For these three tasks, if the participant failed to respond within this time, the event concluded, and the next event appeared. If the participant responded before the time limit, the next event would not occur until the full response window had elapsed. The specific events requiring action, such as which gauge or button needed to be clicked or the number and parameter in the parameter entry task, were randomly generated. Entering an incorrect number or selecting the wrong parameter type in the parameter entry task was counted as an error. For the system and tank monitoring tasks, misclicks were not recorded because the cues were more distinguishable, required only one click, and misclicks had not been observed in the pilot test. For the shooting task, a resting gap with a fixed time was imposed after the maximum allowed response time. Additionally, 10% of the targets did not require a response; shooting these targets was considered an error.

In the high task load conditions, system monitoring and tank monitoring allowed 2 seconds per response, while parameter entry allowed 6 seconds per response since it would require two more clicks than the previous two tasks. This resulted in 18 required responses for each of these three tasks. Each shooting target had 2 seconds for a response, with a new target appearing every 3 seconds, leading to 60 targets in total, of which 54 required a response.

In the low task load conditions, system monitoring and tank monitoring allowed 4 seconds per response, while parameter entry allowed 12 seconds per response since it would require two more clicks than the previous two tasks. This resulted in 9 required responses for each of these three tasks. Each shooting target had 4 seconds for a response, with a new target appearing every 6 seconds, leading to 30 targets in total, of which 27 required a response.

4.5 Statistical Analysis

The analysis explored the effects of four display configurations and two task load levels, resulting in a 4×2 factorial design for the linear mixed models for two performance metrics: response time, and number of errors for each task. Response time was measured from the onset of an event until the participant responded using the mouse or keyboard, with only correct responses included in the analysis. Incorrect responses were analyzed separately as errors because they are not directly comparable in effort or intent, and their heterogeneous causes would inflate variance and confound interpretation of the results. The error count represented the number of missed events or events not completed within the allowed response time. For the parameter entry task, errors also included cases where an incorrect parameter or number was selected. For the shooting task, errors additionally included responses to targets that did not require shooting. The linear mixed models were implemented using the lme4 package [5, 56] in R. Response time was fitted with a Gaussian distribution model, while number of errors, being discrete, was fitted with a Poisson distribution model. In the case of overdispersion, a negative binomial distribution model was used. The models were fitted using full maximum likelihood estimation following a stepwise forward procedure, where a model first starts with the random intercepts for participants then fixed effects are added one at a time for the two predictor variables and their interaction. Likelihood ratio tests were then used to determine if adding each factor significantly improved the model fit. The p-values for these tests were adjusted using the Benjami-Hochberg procedure [6], a balanced approach to control family-wise error rate upon performing multiple comparisons. Significant effects were further explored using Tukey post-hoc comparisons. Residual QQ plots were visually inspected for assumption violations.

The NASA-TLX ratings were not separated by tasks given they reflected overall workload across tasks. The effects of display configuration and task load on each of the six dimensions of the NASA-TLX were analyzed with a two-way repeated-measure analysis of variance (ANOVA). Greenhouse-Geisser corrected p-values were reported where violations of sphericity were detected. Post-hoc pairwise comparisons were conducted using Bonferroni-adjusted p-values to examine significant main effects of display configuration and task load, as well as simple effects of display configuration under each load level in the presence of significant interactions. A two-tailed alpha level of 0.05 was used to determine statistical significance.

5 RESULTS

5.1 Effects of display configurations

5.1.1 Response Time

The display configurations had a significant main effect on *response time* for the parameter entry and shooting tasks, but not for the system monitoring or tank monitoring tasks. Statistical results are presented in Tab. 1. In the parameter entry task, compared to the desktop configuration, the misaligned (*mean difference* $= 1.08$ s, $t = 8.29$, $p < .001$), distant (*mean difference* $= 0.83$ s, $t = 9.25$, $p < .001$), and overlay (*mean difference* $= 1.21$ s, $t = 9.25$, $p < .001$) configurations all resulted in significantly longer response times. Moreover, the overlay configuration led to a significantly longer response time than the distant configuration (*mean difference* $= 0.39$ s, $t = 2.97$, $p = .018$). Similarly, in the shooting task, response times were significantly longer in the misaligned (*mean difference* $= 0.47$ s, $t = 8.24$, $p < .001$), distant (*mean difference* $= 0.44$ s, $t = 7.73$, $p < .001$), and overlay (*mean difference* $= 0.34$ s, $t = 5.88$, $p < .001$) configurations compared to the desktop configuration. Notably, in the shooting task, the average response times in the distant and misaligned configurations were longer than in the overlay configuration, whereas in the other three tasks, they were shorter. The results are shown in fig. 6 (blue bars).

5.1.2 Number of errors

Similar to the response time, the display configurations had a significant main effect on *number of errors* in the parameter entry and shooting tasks, but not for the system monitoring or tank monitoring tasks. Statistical results are presented in Tab. 1. In the parameter entry task, compared to the desktop configuration, the misaligned (*mean difference* $= 1.88$, $t = 3.54$, $p = .003$), distant (*mean difference* $= 1.58$, $t = 2.97$, $p = .018$), and overlay (*mean difference* $= 2.68$, $t = 5.05$, $p < .001$) configurations all resulted in significantly more errors. In the shooting task, there were significantly more errors made in the misaligned (*mean difference* $= 3.65$, $t = 4.78$, $p < .001$) and distant (*mean difference* $= 3.13$, $t = 4.09$, $p < .001$) configurations compared to the desktop configuration. And consistent with the response time results, the average number of errors in the distant and misaligned configurations was higher than in the overlay configuration for the shooting task, but lower in the other three tasks. The results are shown in fig. 7 (blue bars).

5.1.3 Workload

The display configurations had a significant main effect on five out of the six workload dimensions: mental, physical, performance, effort, and frustration. Statistical results are presented in Tab. 1. The desktop configuration was generally associated with lower workload than the three AR configurations, but no significant differences were found among the AR configurations across any of the workload dimensions.

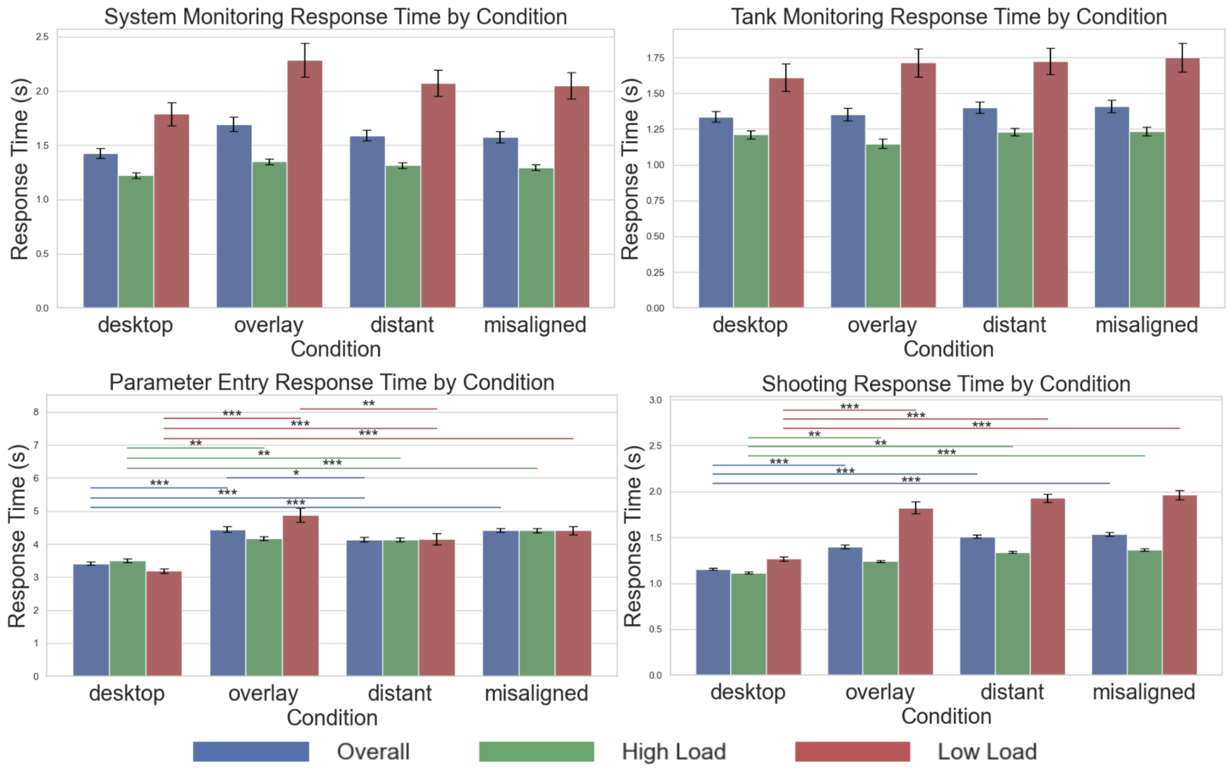

Fig. 6: Mean response times across display configuration conditions with error bars (i.e., standard error). Blue bars show overall performance (i.e., low and high task load), red bars show low task load performance, and green bars show high task load performance. Pairwise brackets for the low/high task load series are shown only when the Condition × Task load interaction is statistically significant. * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

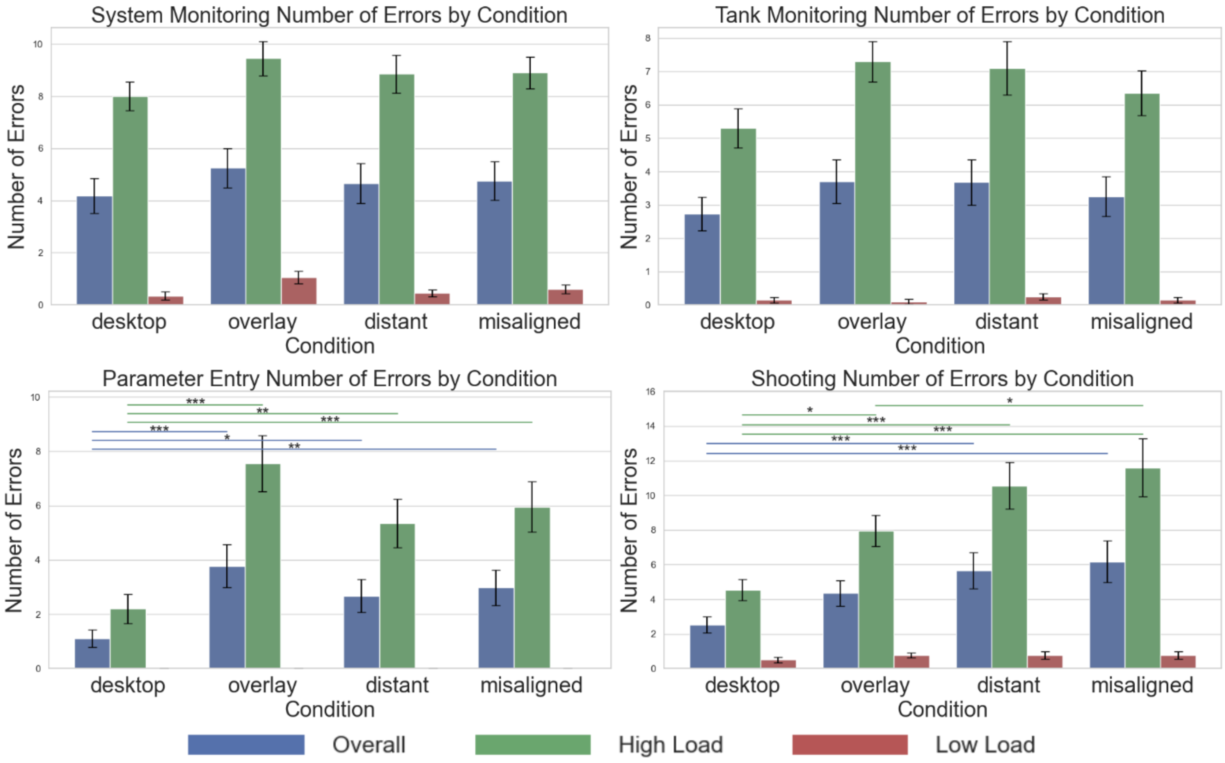

Fig. 7: Mean number of errors made across display configuration conditions with error bars (i.e., standard error). Blue bars show overall performance (i.e., low and high task load), red bars show low task load performance, and green bars show high task load performance. Pairwise brackets for the low/high task load series are shown only when the Condition × Task load interaction is statistically significant. * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

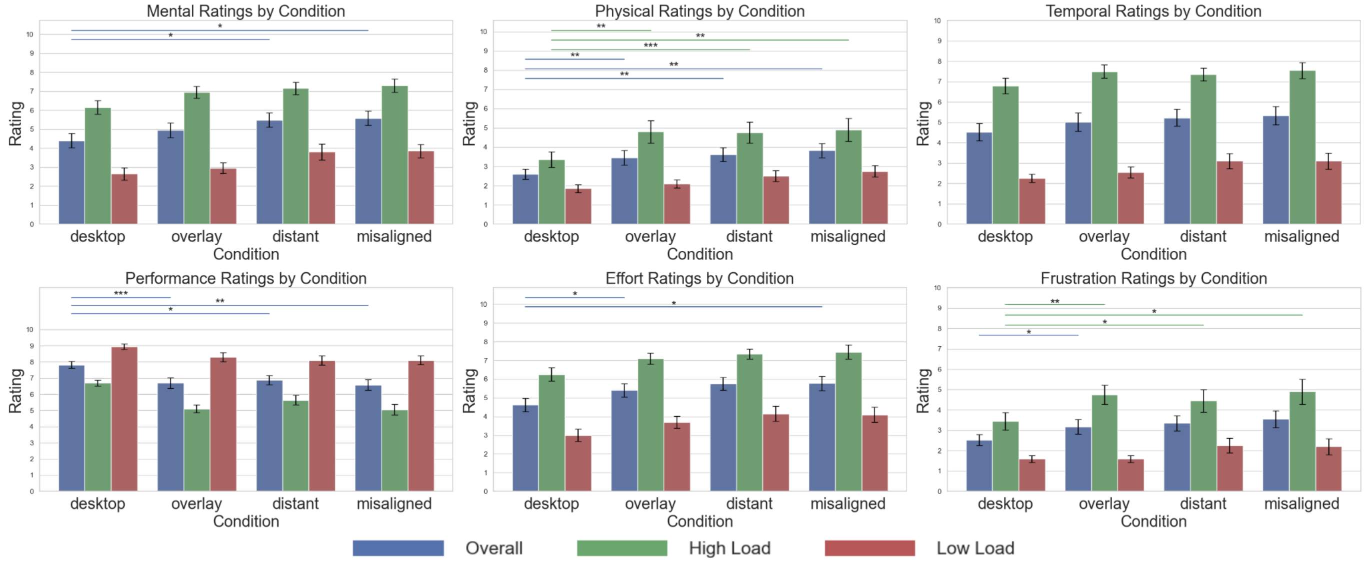

Fig. 8: Mean ratings of the six NASA-TLX dimensions across display configuration conditions with error bars (i.e., standard error). Blue bars show overall ratings (i.e., low and high task load), red bars show low task load ratings, and green bars show high task load ratings. Pairwise brackets for the low/high task load series are shown only when the Condition × Task load interaction is statistically significant. * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

For the *mental* dimension, the distant (*mean difference* $= 1.07$, $t = 3.31$, $p = .022$) and misaligned (*mean difference* $= 1.18$, $t = 3.06$, $p = .039$) configurations had significantly higher ratings than the desktop configuration. For the *physical* dimension, the overlay (*mean difference* $= 0.85$, $t = 3.90$, $p = .006$), distant (*mean difference* $= 1.02$, $t = 4.02$, $p = .004$) and misaligned (*mean difference* $= 1.23$, $t = 3.70$, $p = .009$) configurations had significantly higher ratings than the desktop configuration. For the *performance* dimension, the desktop configuration had significantly higher rating than the overlay (*mean difference* $= 1.12$, $t = 6.63$, $p < .001$), distant (*mean difference* $= 0.95$, $t = 3.53$, $p = .013$) and misaligned (*mean difference* $= 1.25$, $t = 4.63$, $p = .001$) configurations. For the *effort* dimension, the overlay (*mean difference* $= 0.78$, $t = 3.36$, $p = .020$) and misaligned (*mean difference* $= 1.15$, $t = 2.97$, $p = .047$) configurations had significantly higher ratings than the desktop configuration. And for the *frustration* dimension, only the overlay (*mean difference* $= 0.65$, $t = 3.21$, $p = .027$) configuration had significantly higher ratings than the desktop configuration. The results are shown in fig. 8 (blue bars).

5.2 Task load as a moderator of display configuration effects

Compared to low task load, high task load significantly reduced response times in the system monitoring, tank monitoring, and shooting tasks, but had no significant effect on the parameter entry task. A significant interaction between load and display configuration was observed in both the parameter entry and shooting tasks, indicating that the response time differences between display configurations were less pronounced under high task load. As a result, the response time difference between the overlay and distant configurations in the parameter entry task became insignificant under high task load. The results are shown in fig. 6 (red and green bars), and statistical results are presented in Tab. 1.

Compared to low task load, high task load significantly increased the number of errors across all four tasks: system monitoring, tank monitoring, parameter entry, and shooting. A significant interaction between task load and display configuration was observed in the parameter entry and shooting tasks, showing that the differences in error rates between display configurations were more pronounced under high task load. All pairwise comparisons between display configurations for these two tasks became insignificant under low task load. The results are shown in fig. 6 (red and green bars), and statistical results are presented in Tab. 1.

Compared to low task load, high task load significantly elevated workload in all six dimensions. Significant interactions between display configuration and task load were observed in only the physical and frustration dimensions. The interaction effect indicated that differences in ratings of these two dimensions between the desktop and the three AR configurations reached significance only under high workload. The results are shown in fig. 6 (red and green bars), and statistical results are presented in Tab. 1.

6 DISCUSSION

Overall, multitasking performance declined and workload increased when tasks were distributed across AR displays compared to using just a desktop display. The most consistent effects were observed in the parameter entry and shooting tasks, which showed longer response times and more errors under the AR conditions. Within the AR conditions, adding spatial discontinuity between displays generally did not worsen performance compared to the AR condition without spatial discontinuity. Task load amplified the number of errors made and perceived workload differences between display configurations. These findings suggest that both AR use and a higher task load introduces measurable costs to multitasking efficiency and user experience.

6.1 R.Q.1: How does spatial discontinuity between displays affect task performance and perceived workload in multitasking AR?

When comparing the desktop baseline to the AR overlay configuration, the use of AR significantly lowered performance in the parameter entry and shooting tasks. Performance for both these tasks were also lower with the other two AR configurations compared to the desktop baseline. This is reflected in longer response times and higher number of errors made. Although Arefin et al. [2] found that changes in focal plane switching can impair performance with binocular displays, they did not observe significant effects from context switching between virtual and physical environments, either as a main effect or in interaction. Our findings suggest that the combination of context and focal plane switching—even with a relatively small focal distance mismatch (0.76 m)—was sufficient to cause a measurable performance decrement across all AR conditions, regardless of whether there was spatial discontinuity present. This effect may be due to increased eye fatigue (associated with both switching types [2, 19]) and the higher

Table 1: Summary of the main effects of display configuration condition and task load on response time (RT), number of errors, NASA-TLX dimensions, and Condition x Task load interaction. Only statistically significant interactions were included. Statistically significant effects are highlighted in red.

Metric	Task / Dimension	Display config main effect	Task load main effect	Display config × load
RT	System monitoring	$\chi^2(3) = 6.90$, $p = .075$	$\chi^2(1) = 109.74$, $p < .001$	—
RT	Tank monitoring	$\chi^2(3) = 2.49$, $p = .477$	$\chi^2(1) = 9.24$, $p = .002$	—
RT	Parameter entry	$\chi^2(3) = 74.31$, $p < .001$	$\chi^2(1) = 0.004$, $p = .947$	$\chi^2(3) = 15.19$, $p = .002$
RT	Shooting	$\chi^2(3) = 132.93$, $p < .001$	$\chi^2(1) = 89.68$, $p < .001$	$\chi^2(3) = 21$, $p < .001$
# Errors	System monitoring	$\chi^2(3) = 1.11$, $p = .774$	$\chi^2(1) = 275.96$, $p < .001$	—
# Errors	Tank monitoring	$\chi^2(3) = 1.71$, $p = .635$	$\chi^2(1) = 195.51$, $p < .001$	—
# Errors	Parameter entry	$\chi^2(3) = 10.16$, $p = .017$	$\chi^2(1) = 112.57$, $p < .001$	$\chi^2(3) = 25.78$, $p < .001$
# Errors	Shooting	$\chi^2(3) = 9.80$, $p = .020$	$\chi^2(1) = 122.07$, $p < .001$	$\chi^2(3) = 22.91$, $p < .001$
TLX	Mental	$F(3,57) = 6.28$, $p < .001$, $\eta_g^2 = 0.25$	$F(1,19) = 174.15$, $p < .001$, $\eta_g^2 = 0.90$	—
TLX	Physical	$F(3,57) = 9.31$, $p < .001$, $\eta_g^2 = 0.33$	$F(1,19) = 26.39$, $p < .001$, $\eta_g^2 = 0.58$	$F(3,57) = 3.52$, $p = .021$, $\eta_g^2 = 0.16$
TLX	Temporal	$F(3,57) = 2.08$, $p = .112$, $\eta_g^2 = 0.10$	$F(1,19) = 266$, $p < .001$, $\eta_g^2 = 0.93$	—
TLX	Performance	$F(3,57) = 8.40$, $p < .001$, $\eta_g^2 = 0.31$	$F(1,19) = 206.7$, $p < .001$, $\eta_g^2 = 0.92$	—
TLX	Effort	$F(3,57) = 5.60$, $p = .002$, $\eta_g^2 = 0.23$	$F(1,19) = 119.69$, $p < .001$, $\eta_g^2 = 0.86$	—
TLX	Frustration	$F(3,57) = 4.42$, $p = .007$, $\eta_g^2 = 0.19$	$F(1,19) = 30.19$, $p < .001$, $\eta_g^2 = 0.61$	$F(3,57) = 4.99$, $p = .004$, $\eta_g^2 = 0.21$

frequency of display switching in our multitasking setup compared to the visual search tasks used in earlier work [2]. Furthermore, the task itself may have affected performance as well. The present study showed that there were significant effects of AR use on two of the four tasks, i.e., the shooting and parameter entry tasks. This may be due to the fact that the shooting task required more frequent responses and thus was more aversely affected. For the parameter entry task, this typically took longer for participants to complete. As a result, this could have increased the likelihood that participants may have had to continue to switch between displays in order to monitor and keep up with the shooting tasks while completing a parameter entry.

The NASA-TLX perceived workload ratings aligned with the performance results we observed. AR use increased the mental, physical, effort, and frustration dimensions but had no effect on the temporal dimension. The higher physical workload ratings across the three AR conditions may reflect additional strain associated with wearing the headset, such as possible head and neck fatigue [28] or dizziness [51], though these factors were not directly measured in our study. Taken together, these findings support expectation #1: compared to the desktop baseline, AR configurations adversely affected performance and increased perceived workload.

When depth discontinuity was introduced in the distant configuration, there were no significant performance differences compared to the overlay configuration, with one exception: response time in the parameter entry task was faster in the distant configuration. In contrast, for the shooting task, the number of errors made was not significantly different from the desktop baseline configuration in the overlay configuration, but became significantly higher in the distant configuration. One possible explanation for these diverging effects is that the parameter entry task remained fixed in the AR display, while the shooting task was pushed farther away on the physical screen. This could have led participants to prioritize the tasks that were closer in proximity to them. This interpretation is consistent with prior findings on near-space advantages in tasks requiring quick detection and response in AR [9]. Another possible factor is the availability of ambient visual resources. Without spatial separation, motion and color cues from the shooting task may have been detectable peripherally, but as the physical display was moved farther away, these cues likely became less accessible. Prior work has shown that greater focal plane switching distances caused by depth separation can reduce performance in visual search tasks [2], where performance was measured across two displays. In contrast, our multitasking setup assigned different tasks to different displays, allowing us to evaluate them separately. This extends earlier findings by suggesting that depth separation may affect tasks differently depending on whether they are placed closer to or farther from the user. Finally, we suspect that increasing the depth separation beyond the relatively small offset used here could further exacerbate performance costs, especially for tasks that are positioned farther away from the user.

When comparing the misaligned and distant configurations, the lateral visual field discontinuity introduced in the misaligned configuration did not lead to performance difference for any of the tasks. This is consistent with prior work that showed there are no significant effects on response time or the number of errors made when increasing the visual separation angle on either physical or virtual displays [53]. Our results extend these findings to a multitasking environment that combined physical and virtual displays, and further examined visual field separation on top of depth separation. In our study, participants across all configurations could always still see the tasks on both displays at the same time. If the lateral misalignment were increased further so that both displays could no longer be seen simultaneously, a situation made more likely by the limited field of view of most AR headsets, performance could have been affected more. Perceived workload was also not significantly affected by either type of spatial separation. Overall, expectation #2, which predicted that spatial discontinuity between displays would impair performance and increase workload in AR, was not supported based on the findings here.

6.2 R.Q.2: How does task load (low vs. high) modulate these effects?

Higher task load overall adversely affected error rates and increased perceived workload, but decreased response times. The increase in errors made in AR conditions was only significant under high task load as this likely taxed the participants' perceptual capacity [32]. The increase in frustration and physical dimensions of perceived workload also was only significant under high task load. This suggests that challenges introduced by AR——such as context and focal plane switching or potential physical discomfort——may need to be considered more carefully in high-load contexts. For response time, however, higher task load reduced the differences between display configurations: the negative effects of depth discontinuity on parameter entry response time disappeared, and the overall impact of AR was mitigated. This may reflect humans' ability to enhance attentional focus through improved cognitive control under high time pressure [33]. Overall, expectation #3 regarding task load was only partially supported by higher number of errors made and increased perceived workload.

7 DESIGN IMPLICATIONS

Drawing from our findings, we propose four design guidelines for developing AR interfaces that require multitasking across physical and virtual displays.

- **D1. Minimize the needs of frequent switching between physical and virtual displays.** When designers believe that leveraging AR's benefits to support multitasking is truly necessary for their situation, they should be mindful of the costs introduced by context and focal plane switching. Our study showed that even a relatively small focal plane difference, combined with the need to switch contexts, could impair performance in multitasking environments that required frequent switching between display types. Designers should also recognize that AR use was generally associated with higher perceived workload, including increased physical workload potentially caused by headset discomfort. Where possible, related information should be grouped within the same display environment to reduce unnecessary switching.

- **D2. Small spatial discontinuity between physical and virtual displays in AR is acceptable.** If AR use is necessary, our findings suggest that small spatial discontinuities in AR are generally acceptable in multitasking scenarios. In particular, a 1.5 m depth separation and a lateral visual field offset that still allowed both displays to remain within view had minimal impact on performance and perceived workload. This result can give designers more confidence when using the flexibility of spatial placement offered by AR interfaces. However, designers should still be cautious with larger discontinuities.

- **D3. Consider task characteristics when using physical and virtual displays.** Tasks that required more frequent switching or rapid responses were more adversely affected by AR, indicating that minimizing cross-display switching could be more beneficial for these tasks. For cues that may be detected through ambient vision, such as motion and salient color used in this study, designers should be cautious about placing them farther away, as this may reduce their effectiveness. Placing time-sensitive or high-priority tasks closer can also help leverage the near-space advantage, supporting faster detection and response.

- **D4. Account for task load when designing display layouts.** High task load could amplify the negative effects of AR use on errors, frustration, and physical workload. Even when multitasking performance and perceived workload appeared manageable under low task load, designers should anticipate situations where event frequency and time pressure increase. In such cases, display setups could be adapted to ease demands—for example, by simplifying switching paths between the most demanding tasks or reducing the number of simultaneous monitoring requirements.

8 CONCLUSION

In this paper, we investigated the effects of spatial discontinuity between displays in AR on multitasking performance and perceived workload. We also examined how task load modulated these effects. By analyzing response time, number of errors made, and NASA-TLX workload ratings from a simulation across virtual and physical displays and task load conditions, we found that multitasking in AR was associated with increased workload and adversely affected performance. However, among the AR displays, depth and lateral visual field discontinuities produced no significant additional costs overall. Finally, higher task load amplified error rates and perceived workload, while reducing response time differences across display configurations.

This study focused on a small range of spatial discontinuities, including a 1.5 m depth separation and a 20° lateral offset, to evaluate the basic effects of discontinuities in multitasking AR. Larger separations may produce stronger impacts, and future work could test more extreme cases to better understand their influence. Eye tracking can also be used in future work to understand users' visual attention between displays and across tasks. We examined a single setup using one OST-HMD (Magic Leap 2) and a single physical monitor. Future studies could expand on this study by investigating a broader range of display configurations and potentially incorporating more than two displays. While our multitasking simulation captured key elements of monitoring and response, real-world scenarios such as industrial operations or emergency response involve additional complexity, and therefore validating these findings in applied contexts could be helpful. Object-anchored AR may also introduce additional factors (e.g., alignment/registration, occlusion between virtual content and real objects, and body movements) that could further affect performance. Building on recent adaptive interface approaches that dynamically place virtual elements by considering interaction modality and environmental affordances [12], future work could also explore how adaptive placement strategies might support multitasking scenarios that require continuous switching between displays. Finally, given previous work has shown that the impact of context switching between physical and virtual environments may vary based on the users' prior task-related experience [16], future work could also examine how training modulates the performance and workload differences between physical/virtual display combination setups.

REFERENCES

[1] C. Acherki, L. Nigay, Q. Roy, and T. Salque. An Evaluation of Spatial Anchoring to position AR Guidance in Arthroscopic Surgery. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pp. 1–17. Association for Computing Machinery, New York, NY, USA, Apr. 2025. doi: 10.1145/3706598.3713350 2

[2] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. Swan. The effect of context switching, focal switching distance, binocular and monocular viewing, and transient focal blur on human performance in optical see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2014–2025, 2022. 1, 2, 3, 7, 8

[3] D. Ariansyah, J. A. Erkoyuncu, I. Eimontaite, T. Johnson, A.-M. Oostveen, S. Fletcher, and S. Sharples. A head mounted augmented reality design practice for maintenance assembly: Toward meeting perceptual and cognitive needs of ar users. *Applied Ergonomics*, 98:103597, 2022. 1

[4] S. Bang, H. Lee, S. Y. Oh, and W. Woo. AReading with Smartphones: Understanding the Trade-offs between Enhanced Legibility and Display Switching Costs in Hybrid AR Interfaces. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pp. 1–20. Association for Computing Machinery, New York, NY, USA, Apr. 2025. doi: 10.1145/3706598.3713879 2

[5] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *Journal of statistical software*, 67:1–48, 2015. 5

[6] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995. 5

[7] X. Bi, S.-H. Bae, and R. Balakrishnan. Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 65–74, 2010. 2

[8] M. W. Boyce, R. H. Thomson, J. K. Cartwright, D. T. Feltner, C. R. Stainrod, J. Flynn, C. Ackermann, J. Emezie, C. R. Amburn, and E. Rovira. Enhancing Military Training Using Extended Reality: A Study of Military Tactics Comprehension. *Frontiers in Virtual Reality*, 3, July 2022. Publisher: Frontiers. doi: 10.3389/frvir.2022.754627 2

[9] N. Britt, H. Haponenko, and H.-j. Sun. Near-space advantage in a simulated 3d environment: an inhibition of return (ior) study. *Journal of Vision*, 23(9):5016–5016, 2023. 8

[10] I. I. M. Butaslac, Y. Fujimoto, T. Sawabe, M. Kanbara, and H. Kato. Systematic Review of Augmented Reality Training Systems. *IEEE Transactions on Visualization and Computer Graphics*, 29(12):5062–5082, Dec. 2023. doi: 10.1109/TVCG.2022.3201120 1

[11] Y. F. Cheng, A. Carden, H. Cho, C. G. Fidalgo, J. Wieland, and D. Lindlbauer. Augmented Reality In-the-Wild: Usage Patterns and Experiences of Working with AR Laptops in Real-World Settings, Feb. 2025. arXiv:2502.14241 [cs]. doi: 10.48550/arXiv.2502.14241 2

[12] Y. F. Cheng, C. Gebhardt, and C. Holz. InteractionAdapt: Interaction-driven Workspace Adaptation for Situated Virtual Reality Environments. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, pp. 1–14. Association for Comput-

ing Machinery, New York, NY, USA, Oct. 2023. doi: 10.1145/3586183.3606717 9

[13] F. Chiossi, Y. Weiss, T. Steinbrecher, C. Mai, and T. Kosch. Mind the Visual Discomfort: Assessing Event-Related Potentials as Indicators for Visual Strain in Head-Mounted Displays. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 150–159, Oct. 2024. ISSN: 2473-0726. doi: 10.1109/ISMAR62088.2024.00029 2

[14] P. J. Cobb and H. Azizbekyan. Experiments with Mixed and Augmented Reality (MR/AR) for Archaeological Data Collection and Use During Fieldwork: Vision for the Future. *Journal of Computer Applications in Archaeology*, 7(1):370–387, Nov. 2024. doi: 10.5334/jcaa.140 2

[15] S. Davari and D. A. Bowman. Towards Context-Aware Adaptation in Extended Reality: A Design Space for XR Interfaces and an Adaptive Placement Strategy, Nov. 2024. arXiv:2411.02607 [cs]. doi: 10.48550/arXiv.2411.02607 2

[16] M. Drouot, N. Le Bigot, J.-L. de Bougrenet, and V. Nourrit. Effect of context and distance switching on visual performances in augmented reality. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 476–477. IEEE, 2021. 2, 9

[17] M. Eckert, J. Volmerg, and C. Friedrich. Augmented Reality in Medicine: Systematic and Bibliographic Review. *JMIR mhealth and uhealth*, 7:e10967, May 2018. doi: 10.2196/10967 2

[18] B. Fröhler, C. Anthes, F. Pointecker, J. Friedl, D. Schwajda, A. Riegler, S. Tripathi, C. Holzmann, M. Brunner, H. Jodlbauer, et al. A survey on cross-virtuality analytics. In *Computer graphics forum*, vol. 41, pp. 465–494. Wiley Online Library, 2022. 1

[19] J. L. Gabbard, D. G. Mehra, and J. E. Swan. Effects of ar display context switching and focal distance switching on human performance. *IEEE transactions on visualization and computer graphics*, 25(6):2228–2241, 2018. 1, 2, 3, 7

[20] Y. Ghasemi, A. Singh, M. Kim, A. Johnson, and H. Jeong. Effects of Head-locked Augmented Reality on User's Performance and Perceived Workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 65(1):1094–1098, Sept. 2021. Publisher: SAGE Publications Inc. doi: 10.1177/1071181321651169 2

[21] J. Gonnermann-Müller and J. M. Krüger. Unlocking augmented reality learning design based on evidence from empirical cognitive load studies—a systematic literature review. *Journal of Computer Assisted Learning*, 41(1):e13095, 2025. 1

[22] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988. 3

[23] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33, 2008. 2

[24] S. Imamov, D. Monzel, and W. S. Lages. Where to display? how interface position affects comfort and task switching time on glanceable interfaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 851–858. IEEE, 2020. 2

[25] M.-E. Jannat, N. Phan, A. Chadha, K. Katsuragawa, and K. Hasan. HAI-AR: Exploring Hand-Anchored Interfaces in Augmented Reality while Walking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pp. 1–18. Association for Computing Machinery, New York, NY, USA, Apr. 2025. doi: 10.1145/3706598.3713858 2

[26] R. Kalawsky, K. Hill, A. W. Stedmon, C. Cook, and A. Young. Experimental research into human cognitive processing in an augmented reality environment for embedded training systems. *Virtual Reality*, 5:39–46, 2000. 1

[27] M. Kaufeld, M. Mundt, S. Forst, and H. Hecht. Optical see-through augmented reality can induce severe motion sickness. *Displays*, 74:102283, Sept. 2022. doi: 10.1016/j.displa.2022.102283 1, 2

[28] E. Kim and G. Shin. User discomfort while using a virtual reality headset as a personal viewing system for text-intensive office tasks. *Ergonomics*, 64(7):891–899, July 2021. doi: 10.1080/00140139.2020.1869320 1, 2, 8

[29] Y. Kim, H. Kim, and Y. O. Kim. Virtual reality and augmented reality in plastic surgery: a review. *Archives of plastic surgery*, 44(03):179–187, 2017. 1

[30] G. Kramida. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics*, 22(7):1912–1931, 2015. 2

[31] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE international symposium on mixed and augmented reality*, pp. 3–12. IEEE, 2010. 2

[32] N. Lavie. Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3):451–468, 1995. Place: US Publisher: American Psychological Association. doi: 10.1037/0096-1523.21.3.451 3, 8

[33] Y. Li, S. Wang, Q. Shan, and X. Xia. Singleton effect decreases under time pressure: An fnirs study. *Brain and Cognition*, 171:106074, 2023. 8

[34] W. Luo, A. Lehmann, H. Widengren, and R. Dachselt. Where Should We Put It? Layout and Placement Strategies of Documents in Augmented Reality for Collaborative Sensemaking. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–16. ACM, New Orleans LA USA, Apr. 2022. doi: 10.1145/3491102.3501946 2

[35] E. M., V. S. S. V. Prasad, M. Hymavathi, and M. V. A. R. Bahubalendruni. Augmented reality guided autonomous assembly system: A novel framework for assembly sequence input validations and creation of virtual content for AR instructions development. *Journal of Manufacturing Systems*, 72:104–121, Feb. 2024. doi: 10.1016/j.jmsy.2023.11.002 2

[36] R. W. MarklinJr., A. M. Toll, E. H. Bauman, J. J. Simmins, J. F. LaDisaJr., and R. Cooper. Do Head-Mounted Augmented Reality Devices Affect Muscle Activity and Eye Strain of Utility Workers Who Do Procedural Work? Studies of Operators and Manhole Workers. *Human Factors*, 64(2):305–323, Mar. 2022. Publisher: SAGE Publications Inc. doi: 10.1177/0018720820943710 2

[37] A. Matviienko, S. Günther, S. Ritzenhofen, and M. Mühlhäuser. Ar sightseeing: comparing information placements at outdoor historical heritage sites using augmented reality. *Proceedings of the ACM on Human-Computer Interaction*, 6(MHCI):1–17, 2022. 1, 2

[38] M. A. Nacenta, R. L. Mandryk, and C. Gutwin. Targeting across displayless space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 777–786, 2008. 2

[39] F. Nenna, M. Zorzi, and L. Gamberini. Augmented Reality as a research tool: investigating cognitive-motor dual-task during outdoor navigation. *International Journal of Human-Computer Studies*, 152:102644, Aug. 2021. doi: 10.1016/j.ijhcs.2021.102644 2

[40] R. Palmarini, J. A. Erkoyuncu, R. Roy, and H. Torabmostaedi. A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing*, 49:215–228, 2018. 1

[41] L. Pavanatto and D. A. Bowman. Virtual Displays for Knowledge Work: Extending or Replacing Physical Monitors for More Flexibility and Screen Space. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 667–669. IEEE, Orlando, FL, USA, Mar. 2024. doi: 10.1109/VRW62533.2024.00131 2

[42] L. Pavanatto, S. Davari, C. Badea, R. Stoakley, and D. Bowman. Virtual monitors vs. physical monitors: an empirical comparison for productivity work. *Frontiers in Virtual Reality*, 4, Dec. 2023. doi: 10.3389/frvir.2023.1215820 2

[43] L. Pavanatto, C. North, D. A. Bowman, C. Badea, and R. Stoakley. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE virtual reality and 3D user interfaces (VR)*, pp. 759–767. IEEE, 2021. 1

[44] L. Plabst, S. Oberdörfer, F. R. Ortega, and F. Niebling. Push the Red Button: Comparing Notification Placement with Augmented and Non-Augmented Tasks in AR. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*, SUI '22, pp. 1–11. Association for Computing Machinery, New York, NY, USA, Dec. 2022. doi: 10.1145/3565970.3567701 2

[45] A. Raikwar, L. Plabst, A. U. Batmaz, F. Niebling, and F. R. Ortega. Ping! Your Food is Ready: Comparing Different Notification Techniques in 3D AR Cooking Environment. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1157–1166, Oct. 2024. ISSN: 2473-0726. doi: 10.1109/ISMAR62088.2024.00132 2

[46] J. Rasch, M. Wilhalm, F. Müller, and F. Chiossi. AR You on Track? Investigating Effects of Augmented Reality Anchoring on Dual-Task Performance While Walking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pp. 1–21. Association for Computing Machinery, New York, NY, USA, Apr. 2025. doi: 10.1145/3706598.3714258 2

[47] U. Rashid, M. A. Nacenta, and A. Quigley. Factors influencing visual attention switch in multi-display user interfaces: A survey. In *Proceedings of the 2012 international symposium on pervasive displays*, pp. 1–6, 2012. 1, 2

[48] P. Reipschlager, T. Flemisch, and R. Dachselt. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE*

Transactions on Visualization and Computer Graphics, 27(2):1182–1192, Feb. 2021. doi: 10.1109/TVCG.2020.3030460 2

[49] Y. Santiago-Espada, R. R. Myer, K. A. Latorella, and J. R. Comstock Jr. The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide. Technical report, 2011. 3

[50] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, 11(8):11, 2011. doi: 10.1167/11.8.11 2

[51] A. D. Souchet, D. Lourdeaux, A. Pagani, and L. Rebenitsch. A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload. *Virtual Reality*, 27(1):19–50, Mar. 2023. doi: 10.1007/s10055-022-00672-0 2, 8

[52] A. Vovk, F. Wild, W. Guest, and T. Kuula. Simulator Sickness in Augmented Reality Training Using the Microsoft HoloLens. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 1–9. Association for Computing Machinery, New York, NY, USA, Apr. 2018. doi: 10.1145/3173574.3173783 1, 2

[53] A. C. Warden, C. D. Wickens, B. A. Clegg, D. Rehberg, and F. R. Ortega. Information Access Effort: The Role of Head Movements for Information Presented at Increasing Eccentricity on Flat Panel and Head-Mounted Displays. *Human Factors*, 66(8):2057–2081, Aug. 2024. Publisher: SAGE Publications Inc. doi: 10.1177/00187208231204567 2, 8

[54] C. D. Wickens. Multiple Resources and Mental Workload. *Human Factors*, 50(3):449–455, June 2008. Publisher: SAGE Publications Inc. doi: 10.1518/001872008X288394 2

[55] M. Winkler and M. Soleimani. A review of augmented reality heads up display in vehicles: Effectiveness, application, and safety. *International Journal of Human–Computer Interaction*, pp. 1–16, 2025. 1

[56] B. Winter. Linear models and linear mixed effects models in r with linguistic applications. *arXiv preprint arXiv:1308.5499*, 2013. 5

[57] M. Wu, X. Bao, Z. Gu, Y. Tian, Q. Wang, D. Wang, and W. Gao. Optimal information position in OST-HMDs: The effect of eccentricity and direction across two types of AR information. *International Journal of Human-Computer Studies*, 202:103529, Aug. 2025. doi: 10.1016/j.ijhcs.2025.103529 2

[58] X.-D. Yang, E. Mak, D. McCallum, P. Irani, X. Cao, and S. Izadi. Lensmouse: augmenting the mouse with an interactive touch display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2431–2440, 2010. 2

[59] E. Yigitbas, A. Nowosad, and G. Engels. Supporting Construction and Architectural Visualization Through BIM and AR/VR: A Systematic Literature Review. In J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno, and M. Winckler, eds., *Human-Computer Interaction – INTERACT 2023*, pp. 145–166. Springer Nature Switzerland, Cham, 2023. doi: 10.1007/978-3-031-42283-6_8 2

[60] Y. Zhang, K. Chen, and Q. Sun. Toward Optimized VR/AR Ergonomics: Modeling and Predicting User Neck Muscle Contraction. In *ACM SIGGRAPH 2023 Conference Proceedings*, SIGGRAPH '23, pp. 1–12. Association for Computing Machinery, New York, NY, USA, July 2023. doi: 10.1145/3588432.3591495 2