

2D and 3D Augmented Reality Attention Cueing Comparisons in 3D Target Search

Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2024, Vol. 68(1) 1176–1180
Copyright © 2024 Human Factors and Ergonomics Society
DOI: 10.1177/10711813241265077
journals.sagepub.com/home/pro



Brendan Kelley¹, Christopher Wickens¹, Amelia C. Warden¹, Benjamin Clegg², and Francisco Ortega¹

Abstract

This study compared three attentional cueing types to assist the user/participant in visually searching for a target within a simulated naturalistic scene. Participants had to locate a hostile combatant appearing in one of several building windows on an AR-simulated building. Other non-target windows either contain no people or people dressed in non-hostile clothes. We compare the effectiveness of three commonly used attention cues placed in simulated augmented reality on a head mounted display (Magic Leap 2): a 2D wedge arrow pointing toward the target, a 3D Arrow pointing toward the target, and a 3D gaze guidance line, connecting the center of the momentary field of view to the target. Results revealed that the gaze guidance line significantly supported the most rapid cueing and the 3D arrow was slowest and least accurate. The gaze guidance line also induced the fewest head movements. This pattern of results is explained via three mechanisms: the ambiguity of depth perception and the complexity of the 3D arrow, and the ability of the gaze guidance line to always indicate the direction of the target, even when the latter is not in the initial field of view.

Keywords

attention, visual search, augmented reality, display design

Introduction

Searching for and locating targets in a complex real-world wide visual field scene or “search field”; is challenging. Vast amounts of research using 2D flat panel displays have explored this visual search paradigm and established the extent to which it can often be inefficient and time-consuming (Wickens McCarley et al., 2022; Wolfe, 2021), as well as often error-prone (Wolfe et al., 2005). Extension of the paradigm to wide field of view (FOV) 3D scenes is less frequent. However, given the addition of head movements necessary to scan beyond the constraints of a typical computer screen (Poole et al., 2023) and the fact that both peripheral targets and non-targets, sometimes now far out in peripheral vision, will be far less visible, search performance will be expected to degrade accordingly.

It is also well established that some form of attention cueing orienting the searcher’s attention toward what is inferred to be the target, can be beneficial (Raikwar et al., 2024; Warden et al., 2023; Yeh & Wickens, 2001; see Wickens et al., 2022 for a summary). The larger the search field, the greater should be the predicted benefit. In particular, within a wide FOV, the use of a head-mounted display (HMD) to host the cue, enables attention to be directed even when the potential target is far outside the momentary FOV (Gruenefeld

et al., 2018; Yu et al., 2020). Some research, particularly in the human-computer interaction area has examined the optimum kind (physical rendering) of cue (e.g., Gruenefeld et al., 2018; Yeh et al., 2003), but this is typically performed in a target “harvesting” type of paradigms where the goal is to find all objects in a 3D scene rather than search for just one, among a randomly placed set of non-targets, the goal of the current research. In one such study in our laboratory examining the search for a single target we established that, on a head-mounted display, an augmented reality cue pointing toward the target was superior to a cue identifying the target on a minimap (Warden et al., 2023); but this study did not contrast varying AR cues.

In our current research, we chose to compare three HMD augmented reality cueing types to compare their efficiency. Two of the cue types are “2D” in that they register on the HMD display at the same, constant perceptual depth plane.

¹Colorado State University, Fort Collins, CO, USA

²Montana State University, Bozeman, MT, USA

Corresponding Author:

Amelia C. Warden, Colorado State University, 1876 Campus Delivery, Fort Collins, CO 80523-1019, USA.

Email: amelia.warden@colostate.edu

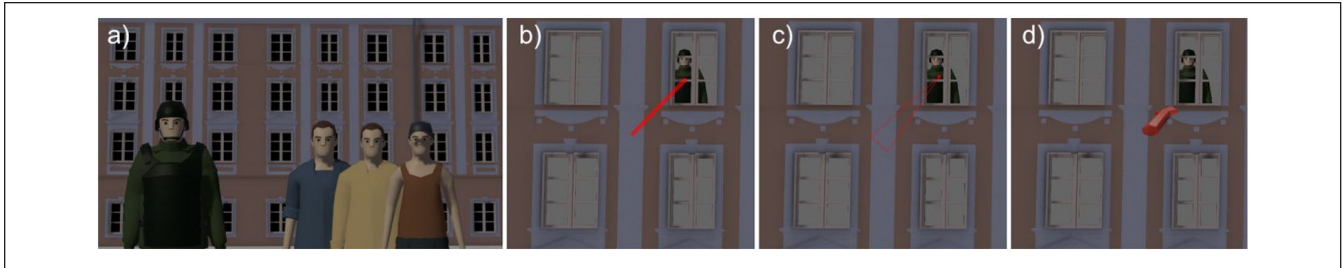


Figure 1. (a) searched objects shown outside the windows, left is target, the other three are non targets, (b) Gaze guidance line pointing to object within the window, as seen in the actual experiment. The end of the guidance line always terminates on the user's viewpoint and hence is visible when the target is not the field of view. (c) 2dWedge, (d) 3d arrow.

The gaze guidance line (GGL) draws a direct vector linkage between the center of the users' current FOV and the cued target. If the latter is out of the current FOV, the GGL terminates at the periphery of the display in a location where it can be extrapolated toward the target and becomes exposed as the head rotates in that direction (Mifsud et al., 2022). The 2D wedge cue provides similar functionality. It is composed of a triangle with the tip overlaid onto the target location transposed from the 3D space onto a 2D user interface plane that scales with the distance from the target. The 2D wedge creates less clutter on the display, a factor well established to inhibit visual search (Warden et al., 2023). In contrast, the 3D arrow cue examined frequently in computer science research (e.g., Gustafson et al., 2008; Wieland et al., 2002) contains perspective, in pointing in depth, toward the target location. Thus, the arrow changes in shape as the head rotates, becoming more elongated as the head moves away, and more foreshortened as the orientation approaches the target. On the one hand, this makes the view more "natural," but on the other hand, this rendering suffers from some of the same biases and distortions that have been well documented for 3D perception in depth (Smallman & St. John, 2005; Wickens Vincow & Yeh, 2005); In the current case, these biases include changes (i.e., inconsistency) in the visual shape of the arrow as the head rotates.

This research thus addresses two primary questions: What is the overall effectiveness of HMD attention cueing given the wide FOV search field, and how is this effectiveness moderated by the type (two vs. three dimensionality) of the cue used.

Methods

Participants

Fourteen participants volunteered for the study. Eight male, five female, one non-binary. Their mean age was 22.8 years with all but two reporting current undergraduate status. Six of these individuals reported prior use of an AR headset. All individuals were compensated for their time with either a

\$25 Amazon gift card, or extra credit for a participating course.

Task and Equipment

We employ a laboratory based approach to answering these questions. Participants wore a mixed reality HMD, in which the 3D search field was rendered as the wall of a 4-story building, viewed from the ground level and containing 32 windows (Figure 1). The overall FOV of the wall was 116° (lateral) by 90° (vertical) this was rendered virtually on an AR HMD with an FOV of 55° by 45° . Twenty windows contained people; one person on each search trial, the target, was wearing a soldier's uniform. When the search trial started, the participant was instructed to scan the wall until (or if) the target is located, and so indicate by using a motion tracked controller to point at and select the target via the trigger button. Software assessed if they selected the true target when the button was pressed (correct), or if they selected an empty window or distractor elsewhere (error). On 25% of the trials, the target was uncued and on the 75%, one of the three cues are presented.

Design

The four conditions were presented in blocked, counterbalanced order, with two blocks (replications) for each condition. There was a 20s rest break between each of the eight blocks. The cue conditions were presented in a counterbalanced order within each block.

Results

Data were first examined for outliers using the IQR method with a threshold of 1.5 times the IQR. Only one outlier for accuracy was identified and removed via this method.

Response Time

Figure 2 presents the RT data across the four conditions.

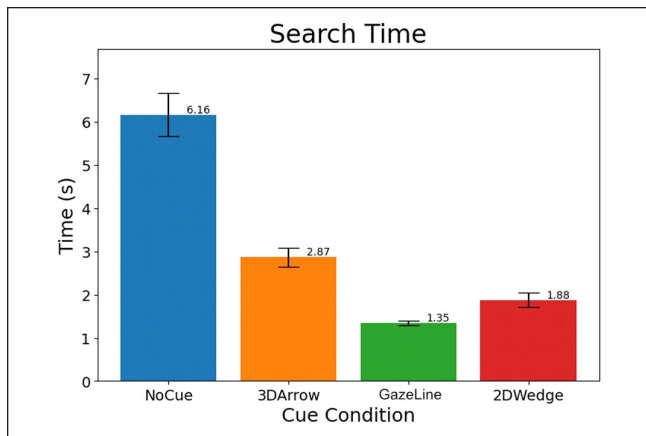


Figure 2. Search time.

Repeated Measures MANOVA showed a main effect of cue condition ($F[3, 42]=56.6; p<.01$) indicating the obvious benefits of all cues ($M=2$ s) over the control condition ($M=6.2$ s).

Within the cue types, paired contrasts revealed that the gaze guidance line supported more rapid responding than the 2D wedge ($t_{(13)}=3.24; p<.01$) and the 2D wedge in turn supported faster responses than the 3D arrow ($t_{(13)}=6.24; p<.01$).

Accuracy

Figure 3 presents the mean accuracy data.

ANOVA on the accuracy data revealed a significant main effect of cue conditions ($F[3, 39]=4.88; p<.01$) revealing that all cues ($M=97\%$) improved accuracy relative to the control condition ($M=87\%$). Within the three cue types, the 3D arrow was less accurate than either the 2D gaze guidance line ($t_{(13)}=2.34; p<.05$) or the 2D wedge ($t_{(13)}=2.69; p<.05$). The 2D wedge and the gaze guidance line did not differ in accuracy ($p>.10$).

Head Movements

Figure 4 presents the data for the mean amount of head movements, assessed by the integration of movement over time. The main effect of condition was highly significant ($F[3, 42]=58.3; p<.01$) with the control condition ($M=2,096$) yielding over twice as much head movement as the mean of the cued conditions ($M=900$). Within the latter, the 3D arrow yielded more head movements than the 2D wedge ($t_{(13)}=2.16; p<.05$) and the gaze guidance line ($t_{(13)}=3.65; p<.01$). The gaze guidance line yielded slightly less head movement than the 2D wedge, at a level approaching conventional statistical significance ($p=.051$).

Discussion

The current results clearly support the demonstrated value of attention cueing, and it is not surprising that directing attention

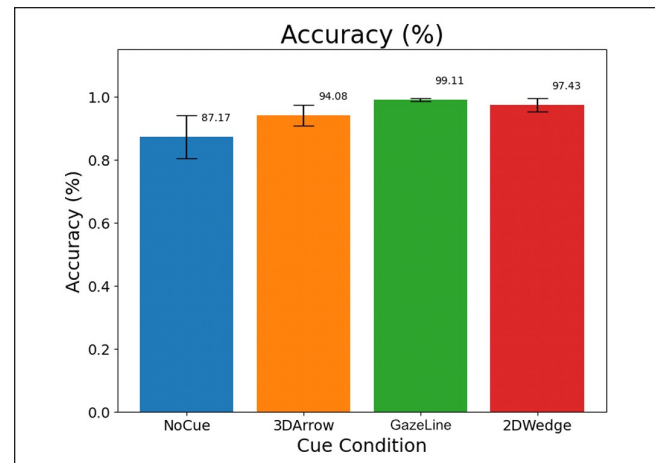


Figure 3. Mean accuracy data.

toward the suspected target would shorten response time. But the *amount* of shortening was nontrivial, from 6 s to approximately 2 s for the two best cues. In real-world circumstances like a combat zone or high-speed driving, shortening by this amount could be of great operational significance. It is also not surprising, in comparing Figures 2 (RT) and 4 (head movement) that the two variables are almost perfectly correlated across conditions, a finding that indicates that the prominent benefit of cueing results from the reduction in time-consuming head movements to search for a cue that may not always be in the initial field of view. This will be true of both of the cues which presented arrows close to the target; but will not impose a penalty for the gaze guidance line which, at the beginning of every trial has one end always in view, and that end leads in a direction that directly signals the direction of head rotation necessary to find the target. This analysis in turn indicates that the head rotation-induced RT penalty probably results primarily on those trials in which a cue was not initially in view, and the participant initially scans in the wrong direction; not finding the target on that side, they then reverse scan to the other side of their forward field of view.

The results reveal that there is substantially less of a cueing benefit for the 3D cue, despite its greater realism. These results converge with many others: the greater realism of “3D,” whether rendered by the perspective, stereo, or any number of other depth cues, while often preferred by a user (Smallman & St. John, 2005) frequently performs less well than its 2D plan-view counterpart, for reasons related to the ambiguity of 3D viewing (Ellis et al., 1991; Wickens et al., 2005). Second, the large cueing benefits observed here were realized with perfect cues that always highlighted the target. In operational circumstances, imperfect machine learning algorithms may need to identify targets within a very short time frame, and that identification could be imperfect. Other research in our laboratory has revealed the costs of that imperfect reliability (Raikwar et al., 2024; Warden et al., 2023; Yeh & Wickens, 2001) and such costs appear to be

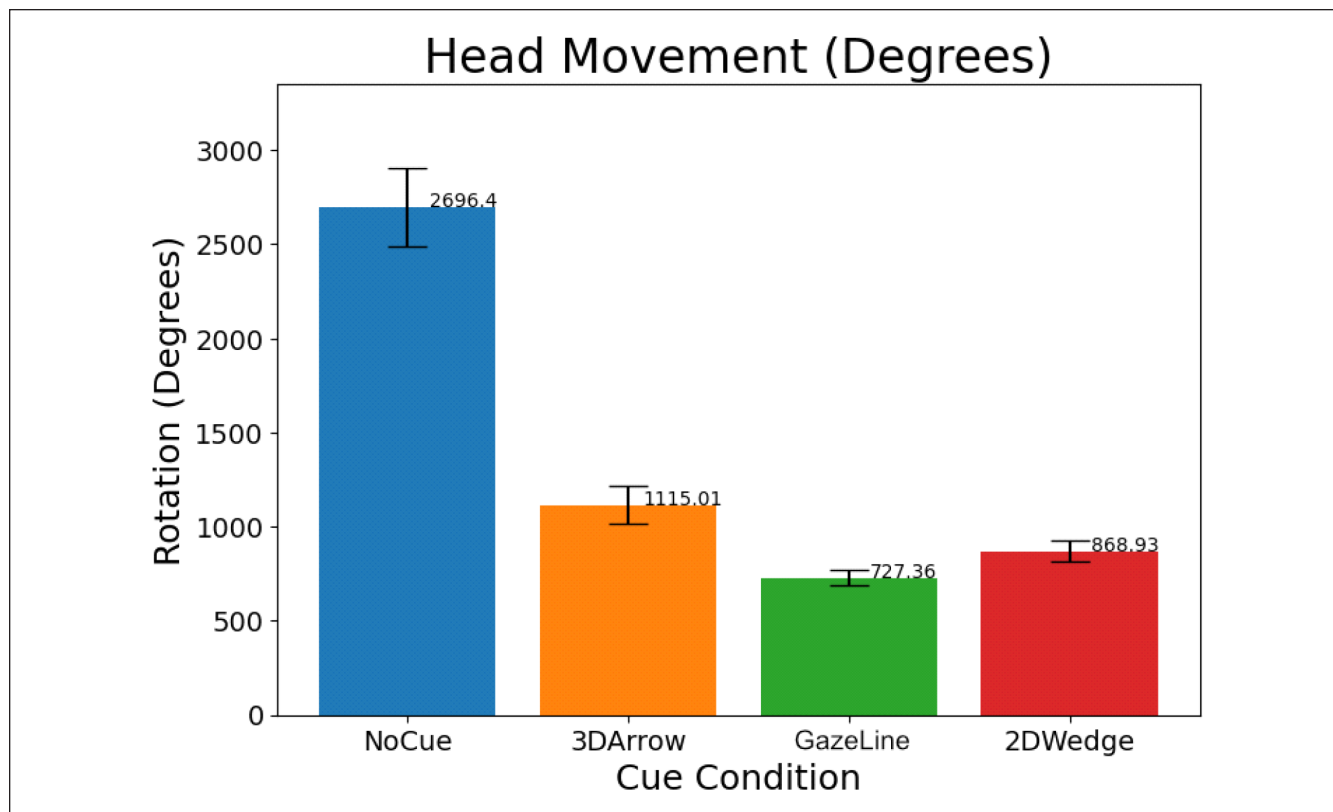


Figure 4. Head movement measured in degrees.

even amplified for those cues that work better when they are correct (Warden et al., 2023).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by grants from the Office of Naval Research # N00012-21-1-2949 and N00014-21-1-2580. Dr Peter Squire was the scientific/technical monitor.

References

- Ellis, S. R., Smith, S. R., Grunwald, A. J., & McGreevy, M. W. (1991). Direction judgement error in computer generated displays and actual scenes. In S. R. Ellis (Ed.), *Pictorial communication in virtual and real environments* (pp. 504–526). Taylor and Francis.
- Gruenefeld, U., Ali, A. E., Boll, S., & Heuten, W. (2018). Beyond Halo and Wedge: Visualizing out-of-view objects on head-mounted virtual and augmented reality devices. *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services* (pp. 1–11). <https://doi.org/10.1145/3229434.3229438>
- Gustafson, S., Baudisch, P., Gutwin, C., & Irani, P. (2008). Wedge: Clutter-free visualization of off-screen locations. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 787–796). <https://doi.org/10.1145/1357054.1357179>
- Mifsud, D., Wickens, C., Maulbeck, M., Crane, P., & Ortega, F. (2022). The effectiveness of gaze guidance lines in supporting JTAC's attention allocation. *Proceedings of 66th annual meeting the human factors and ergonomics society*. Sage Press.
- Poole, C. A., Warden, A. C., Wickens, C. D., Raikwar, A., Clegg, B. A., & Ortega, F. R. (2023, September). Information access costs with a wide-angle desktop display. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 67, No. 1, pp. 947–952). SAGE Publications.
- Raikwar, A., Warden, A., Mifsud, D., Wickens, C., Batmaz, A., Kelley, B., Clegg, B., & Ortega, F. (2024). Beyond the wizard of Oz: Negative effects of imperfect machine learning to examine the impact of reliability of augmented reality cues on visual search performance. *Transactions on Visualization and Computer Graphics*.
- Smallman, H. S., & St. John, M. (2005). Naïve realism: Misplaced faith in the utility of realistic displays. *Ergonomics in Design*, 13, 6–13.
- Warden, A. C., Wickens, C. D., Rehberg, D., Ortega, F. R., & Clegg, B. A. (2023). Fast, accurate, but sometimes too-compelling support: The impact of imperfectly automated cues in an augmented-reality head-mounted display on visual search performance. *IEEE Transactions on Human-Machine Systems*, 53(6), 1061–1072. <https://doi.org/10.1109/THMS.2023.3302152>

- Wickens, C. D., Vincow, M., & Yeh, M. (2005). Design applications of visual spatial thinking: The importance of frame of reference. In A. Miyaki & P. Shah (Eds.), *Handbook of visual spatial thinking*. Oxford University Press.
- Wolfe, J. M. (2021). Guided search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28(4), 1060–1092.
- Wickens, C. D., McCarley, J. M., & Gutzwiller. (2022). *Applied attention theory* (2nd ed.). CRC Press, Taylor & Francis.
- Wieland, J., Garcia, R. C. H., Reiterer, H., & Feuchtner, T. (2022). Arrow, Bézier curve, or Halos?—Comparing 3D out-of-view object visualization techniques for handheld augmented reality. *2022 IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 797–806). <https://doi.org/10.1109/ISMAR55827.2022.00098>
- Wolfe, J., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, 435(7041), 439–440.
- Yeh, M., Merlo, J. L., Wickens, C. D., & Brandenburg, D. L. (2003). Head up vs. head down: The costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors*, 45(3), 390–407.
- Yeh, M., & Wickens, C. D. (2001). Display signaling in augmented reality: Effects of cue reliability & image realism on attention allocation and trust calibration. *Human Factors*, 43(3), 355–365.
- Yu, D., Liang, H.-N., Fan, K., Zhang, H., Fleming, C., & Papangelis, K. (2020). Design and evaluation of visualization techniques of off-screen and occluded targets in virtual reality environments. *IEEE Transactions on Visualization and Computer Graphics*, 26(9), 2762–2774. <https://doi.org/10.1109/TVCG.2019.2905580>