2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 979-8-3315-0691-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISMAR-Adjunct64951.2024.00051

Addressing Human Factors Related to Artificial Intelligence Integrated Visual Cueing

Brendan Kelley* — Colorado State University
Aditya Raikwar† — Colorado State University
Ryan P. McMahan‡ — Virginia Tech
Benjamin A. Clegg§ — Montana State University
Chris D. Wickens¶ — Colorado State University
Francisco R. Ortega‖ — Colorado State University

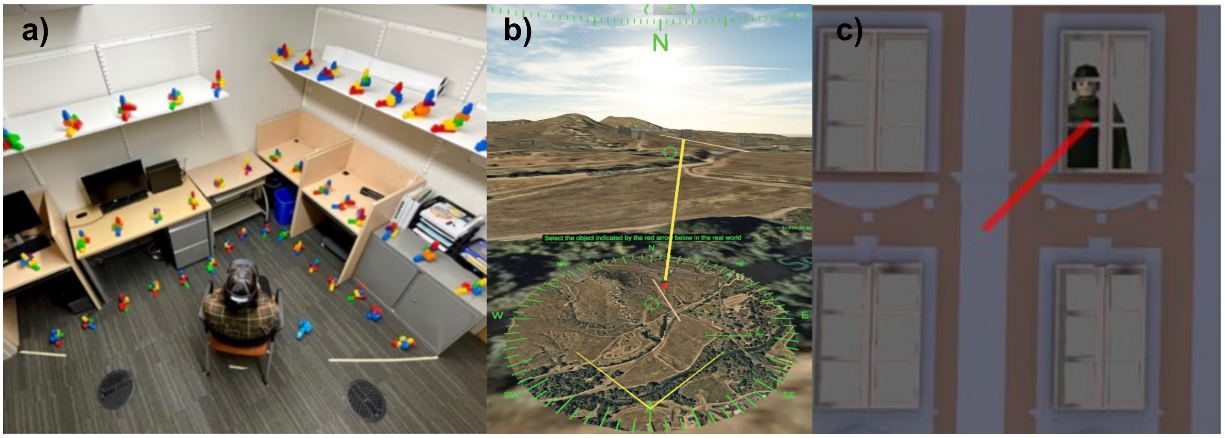

Figure 1: Views from studies conducted related to AI and XR cueing. a) the search environment and task used for Study 3 and Study 4 [17, 18, 23]; b) HMD view of the environment and task used for Study 2 [13], c) HMD view of the visual search task and the gaze line cue used for Study 1 [10, 9].

Abstract

A variety of assistive extended reality (XR) visual cueing techniques have been explored over the years. Many of which provide significant benefits to tasks such as visual search. However, when the cueing system is erroneous, performance may instead suffer. Factors such as automation bias, where an individual trusts the cueing system despite errors in the cueing may affect task efficacy (i.e. completion time, accuracy, etc.). In some cases, such as with automation bias, these hindrances may be the product of artificial intelligence (AI) integration. Despite this, there may be benefits to using adaptive AI-based cueing systems for XR tasks. However, aspects such as the flow of information, automation accuracy, communication of confidence, or the refusal of output must be considered to build effective AI adaptive cueing systems. In this paper, we discuss four studies conducted by our group that explore visual cueing and AI. We then discuss potential future avenues for integrating AI into cueing techniques to minimize automation bias and cognitive demand on users, as well as, improve overall cueing outcomes.

Index Terms: Extended Reality, Virtual Reality, Augmented Reality, Visual Cueing, Machine Learning, Automation, Adaptive Systems

*e-mail: brendan.kelley@colostate.edu
†e-mail: adityaraikwar@gmail.com
‡e-mail: rpm@vt.edu
§e-mail: benjamin.clegg@montana.edu
¶e-mail: chris.wickens@colostate.edu
‖e-mail: f.ortega@colostate.edu

1 Introduction

A variety of factors influence visual cueing for extended reality (XR). Design aspects such as head-mounted displays (HMDs) capabilities, communicated information, salience, cue triggers, task context, etc., affect the efficacy of cueing [4, 8, 7]. In addition to these design aspects, human factors, such as attention [25] and cognitive resources [22] affect cueing intervention efficacy; namely through automation bias [3, 20] and information overload (often via display clutter) [14].

In this paper, we will discuss several of our studies surrounding visual cueing and artificial intelligence (AI). Study 1 (see Fig. 1c) supports the usage of cues for visual search in XR [10, 9]. However, these results were produced with perfect cueing. With these cues, if present, they point to the target location without fail. Study 2 (see Fig. 1b) begins to explore error [13]. When automation (or simulated automation) was added to the cueing system via machine learning (ML), diminished returns surfaced when utilizing cues as demonstrated by Study 3 [23] and Study 4 [17, 18] (see Fig. 1a). Many of the issues that emerged when ML was incorporated can be attributed to automation bias, where the user trusts the automated system, even when the system is erroneous [3, 20]. This issue can even persist when collaborating in groups [21]. Clutter effects can further compound these issues by over-taxing- the limited cognitive resources a user possesses.

Adaptive cueing approaches can be adopted to address these issues, some of which are the product of AI integration. Adaptive cueing uses AI, such as ML, to dynamically drive the cueing system implemented for a given task, based on computer sensing of the environment after training the model for a specific purpose. This approach has the potential to reduce cognitive load dynamically, adapt to user skill levels, and help improve overall cueing outcomes. However, new considerations emerge, such as information flow, automation accuracy, automation confidence, and how to

2771-1110/24/$31.00 ©2024 IEEE
DOI 10.1109/ISMAR-Adjunct64951.2024.00051

handle uncertainty.

2 XR Visual Cueing

The visual search paradigm asks users to find a target from among other distractor objects, from a complex environment, or a combination of both [15, 5]. It is a common task used for XR cueing and has been applied to a variety of use cases such as airborne surveillance [1], finding office supplies [16], or navigating stairs for low-vision individuals [27]. Several studies have explored important aspects such as cue subtlety [24], lighting [11], out-of-view targets [6, 2], and large area searches [12]. Many of these prior studies have found cueing techniques to be beneficial for the assigned task, improving things such as time and accuracy metrics.

Several prior works are in line with these results. When using gaze guidance lines in Study 2 to identify buildings in a virtual environment for a simulated Joint Tactical Air Controller (JTAC) target identification task, Mifsud et al. [13] found the use of gaze guidance lines improved both response time and accuracy. In our Study 1 using visual cues to identify hostile targets in virtual buildings, we compared three cue designs (gaze lines, 2D wedge, and 3D arrow) to a baseline no-cue condition [9, 10]. Our results showed an improvement in both search time and search accuracy when using any cue condition (see Figure 2). However, the greatest reduction in time and increase in accuracy was produced by the gaze line cue [9].

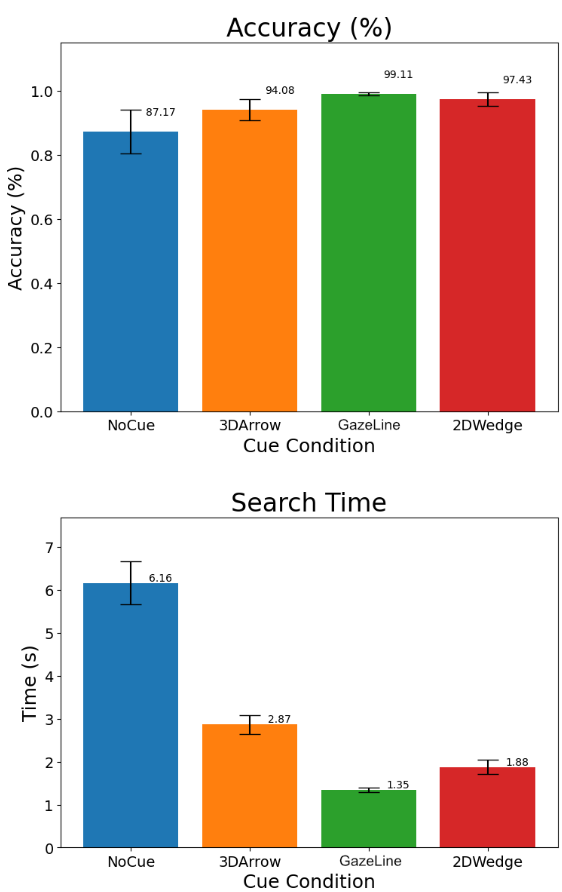

Figure 2: Accuracy and response time measures from Study 1 [9, 10]. Each cue condition improved over the baseline no cue condition.

In addition to the improvements to both search time and accuracy from cue integration, we also found a reduction in the total head rotation exhibited by participants with the cues in Study 1 [10]. When using a cue, participants' total head rotation°was reduced from 2696.4°(about 7.5 full rotations of the head) to between 727.36°and 1115.01°(or about 2 to 3 full rotations of the head) with the gaze line providing for the greatest reduction (see Figure 3. This is attributed to the readily accessible spatial information provided by the cues and the inclusion of both position and direction information in the gaze line and 2D wedge designs (as opposed to the 3D arrow, which only includes direction information). This reduction may also translate into a reduced sense of fatigue when using these systems. However, these benefits are predicated on the perfectly cued nature of these works. In each of these studies, the cue, when present, always pointed to the target location, but the introduction of erroneous cueing has been found to introduce automation bias effects.

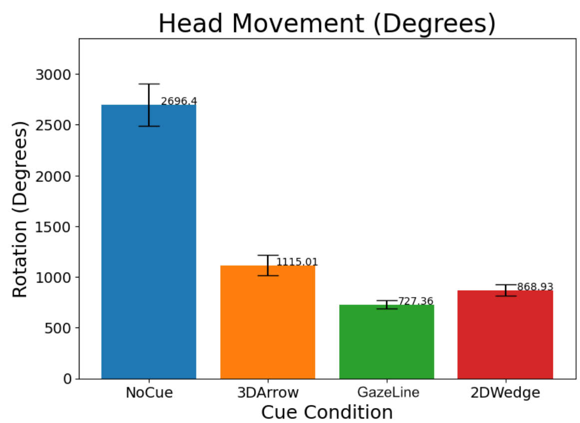

Figure 3: Rotation measures from Study 1 [10]. Each cue condition reduced the total head rotation over the baseline.

3 Automation Bias

Automation bias occurs when the user of an automated system, such as an AI cueing system, trusts the system's decisions even when erroneous [3, 21, 20]. In experiment 2 of Study 2 [13] erroneous cues were introduced into the JTAC target identification task. In 10% of the conditions, the gaze line cue would point to the wrong building. This introduction of error reduced the efficacy of the cueing intervention and appeared to have introduced an automation bias effect.

3.1 Bias with Machine Learning

A follow-up study (Study 3) conducted by our group introduced error by simulating output that would be produced by an ML agent trained on target data [23]. Three cue conditions were tested: 2D arrows, minimap, and icons (in the form of images of the target stimuli). Each cue was presented with either a 17% error rate (to simulate ML-driven output with an 83% accuracy rate) or perfectly cued. Cues boosted search performance even when erroneous (see Figure 4), however the cues did suffer a performance decrement (reduced response time and accuracy) when erroneous. Despite the performance change cueing still provided a benefit; in particular, the arrow cue was still the fastest and most accurate [23]. This performance drop may indicate an automation bias effect where people followed the cue without confirming it's accuracy.

Many prior studies exploring automation bias used an imperfect system similar to our Study 3. [23]. These often used a fixed error rate, however, a true ML-driven cueing system would be subject to variability based on environmental conditions. This motivates our follow-up study to Study 3 [23]. In Study 4 [17, 18], the search

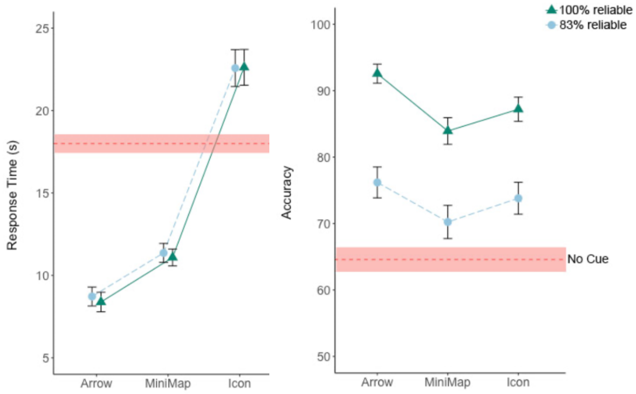

Figure 4: Comparison of cues with different simulated accuracy rate from Study 3 [23]. The arrow cue provided for the fastest response time and highest accuracy even when the cue was 83% accurate. All cues improved over the baseline for both response time and accuracy, withe the exception of the icon cue response time.

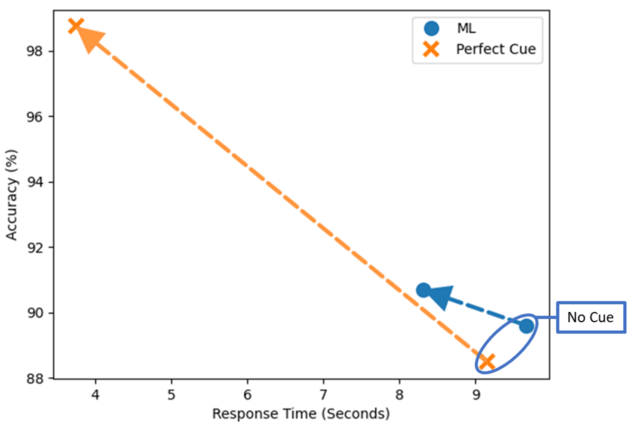

Figure 5: Speed-accuracy trade-off from Study 4 [17]. When the cue perfectly cued the target the highest accuracy and fastest response time was observed. When ML was used to drive the cue the response time increased and the accuracy decreased, however the use of ML cues still improved both accuracy and response time.

task was similar to Study 3 [23], except a single cue (2D arrow) was compared to a no cue condition with either perfect cueing or a YOLOv5 ML model. Imperfection in the cueing system was introduced using this live ML model with an average accuracy of 88.9%. In line with Study 1 and Study 3, the use of perfect cueing provided for the strongest response in regards to both accuracy and response time, however, when the live ML agent was used to drive the cues both accuracy and response time suffered (although cueing was still beneficial). Figure 5 shows the trade-off experienced between speed and accuracy as well as the effects of both cueing condition and perfect vs ML cueing.

A delay was also sometimes observed in the response time for the imperfect cueing condition driven my ML [17]. In these cases the participant did not blindly follow the automated recommendation, but instead noticed the error. This then prompted them to research the space for the correct target. However, if they were then unable to find the correct target, based on their imperfect recall of an image shown at the beginning of the trial, they would typically default to following the erroneous ML decision [17].

These works highlight one of the key challenges with using AI to drive cueing systems: the allure of trusting the AI system even when it is erroneous. A robust adaptive cueing approach may prove beneficial in addressing this issue and other factors that exist when considering the human element of XR cueing.

4 Adaptive Cueing

Adaptive cueing approaches incorporate AI techniques into the cueing system to dynamically adjust the cueing experience. Seeliger et al. [19] applied adaptive cueing to a navigation task. Participants were asked to navigate a room and retrieve items from boxes located throughout. The adaptive system produced comparable results to the always-on system (the format most cueing systems utilize), but with the benefit of higher user preference and reduction of clutter effects. Another recent study explored the use of adaptive-auditory cues for gait control with Parkinson's disease [26]. As with [19], Wu et al. [26] found the use of adaptive cueing techniques to be beneficial, however this study did not present the auditory cues in an XR setting.

5 Implementation of Adaptive Cueing

One of the key challenges with deploying an adaptive cueing system is gathering enough relevant data to train and build AI models (a problem that is not uncommon in the ML community). In the case of adaptive cueing, performance and interaction data are the most suitable measures. This includes performance data, such as response time or accuracy, as well as physical/physiological measures, such as movement, gaze, pupillometry, heartbeat, etc. In the case of [19], context data (i.e., position, distance to target, etc.) and user input data were used to train their model.

Once collected, the goal is to capitalize on features of that dataset that can help predict when a user would benefit from cueing interventions. This may be when their response time is slowed or idled. Another indication of the need for cueing intervention could be when their heart rate increases, or the number of saccades increases without fixations on the target or relative spatial features. Other task specific data may be relevant, especially when the data indicates a potential danger to the user (i.e. overheating machinery, active automated robotic agents, etc.). In any case the goal is to reduce the amount of cognitive effort required to engage with the task. This may in turn allow the user to attend to the target, distractors, and environment to better identify when an error has occurred.

There is also a question of user agency and the overall flow of information with cueing systems (adaptive or otherwise). When considering a use case with a top-down hierarchical structure, such as military or first responder, in addition to the user and the AI agent driving the cueing, another source may be pushing information to the user. This could be dispatch or a command center, a collaborative team of XR users, etc. This becomes a complex problem that could overwhelm the user if proper design considerations are not considered, as there are now many channels of information that may present cued information to the user.

5.1 Reducing Cognitive Load

Attention theory [25] posits that individuals have a limited number of cognitive resources at their disposal. The amount of resources depends on the individual, their environment, the time of day, and a host of other contextual factors. If an individual becomes overloaded and the task at hand requires more cognitive resources than they have, their performance will suffer. Conversely, if a task requires so few cognitive resources as not to engage the individual, this will also affect performance.

In the case of XR cueing, providing too much information to an individual may lead to diminishing returns via cluttering effects [14]. Display clutter can occur with any display but is especially problematic with low FOV augmented reality (AR) displays. To combat this, an adaptive cueing system that dynamically changes the amount of cued information can be implemented. This does, however, need to take into consideration the nature of the cued information, including its importance to the task, the danger posed to the user, temporal relevancy, etc. These aspects become even more difficult to manage when considering the confidence and accuracy of the cueing system.

5.2 Confidence and Output

With ML models, a key feature for determining output is confidence. This metric helps the model to determine what output to present to the user. In most cases, with regard to adaptive XR cueing, this means that whatever target location has the highest confidence rating would be cued. For a majority of tasks, this would be acceptable, even with errors as a factor, but in high-stakes, potentially life-threatening scenarios, such as military or search and rescue operations, these errors may be unacceptable. Unfortunately, it may not be enough to simply tag each cue with a confidence value as this may increase the number of cues (i.e., three locations all produce a .33 confidence score, so all three are cued) and increase clutter by presenting the confidence information alongside the cues. This would, in turn, affect the cognitive load of the user.

Instead of presenting this information it may be pertinent to selectively ignore information depending on the system's confidence, however this may come with other costs, especially when considering the use of adaptive XR in high stakes scenarios. A missed enemy soldier, an inaccurate injured victim location, or some other misrepresented data may quickly become problematic, especially if automation bias is in play. In these situations, opting not to provide information may be the better choice in order to force the user to rely on their own perception for task completion. In addition, there may also be situations where the user must maintain focus and should have the option to dismiss or completely turn off the adaptive cueing system.

5.3 Other Considerations

Collaborative control of the adaptive cueing system can become quite challenging, especially if the system is dependent on a complex flow of information from the environment, the user, collaborative users, and other hierarchical structures. Allowing the user to switch off or dismiss information that they believe to be erroneous or unhelpful can improve this experience, but there may also be times when the adaptive system needs to begin re-cueing without the user's knowledge (i.e. imminent danger in a combat zone or a more critical patient going into cardiac arrest). In these cases, it may be necessary to remove control of the dismissal options from the user. However, this can lead to complications from clutter or cognitive load effects.

For adaptive XR cueing, the best scenario is to develop robust models that have a low error rate. However, this is not always possible, given the nature of AI (particularly ML). As such consideration of the flow of information, how to communicate or act upon confidence, and methods of shifting adaptive system control should be considered. To address each of these design considerations, further exploration is required from both the XR and the AI community.

6 Conclusion

Integration of AI into XR cueing techniques may be inevitable, as the cueing system will need to be able to identify and act upon relevant spatial information in an environment without markers. While cueing systems can provide a benefit to the user, these benefits quickly diminish when the cueing system becomes inaccurate, a problem that is compounded by automation bias and clutter effects. In order to counteract these effects adaptive-cueing systems driven by AI may prove to be a useful tool, however, implementation of such systems requires careful consideration of the factors surrounding XR cueing. In addition to the cue design and human factors, aspects such as the flow and control of information, the nature of the data being utilized for adaptive cueing, the handling of imperfect information and confidence, as well as a host of other factors, require consideration.

References

[1] N. Barbotin, J. Baumeister, A. Cunningham, T. Duval, O. Grisvard, and B. H. Thomas. Evaluating Visual Cues for Future Airborne Surveillance Using Simulated Augmented Reality Displays. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 213–221. IEEE, Christchurch, New Zealand, Mar. 2022. doi: 10.1109/VR51125.2022.00040 2

[2] N. Biswas, A. Singh, and S. Bhattacharya. AroundArrow: Off-Screen POI visualization for handheld Augmented Reality in vertically dense regions. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 570–575. IEEE, Singapore, Singapore, Oct. 2022. doi: 10.1109/ISMAR-Adjunct57072.2022.00119 2

[3] M. Cummings. Automation Bias in Intelligent Time Critical Decision Support Systems. In *AIAA 1st Intelligent Systems Technical Conference*. American Institute of Aeronautics and Astronautics, Chicago, Illinois, Sept. 2004. doi: 10.2514/6.2004-6313 1, 2

[4] K. R. Dillman, T. T. H. Mok, A. Tang, L. Oehlberg, and A. Mitchell. A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, Montreal QC Canada, Apr. 2018. doi: 10.1145/3173574.3173714 1

[5] M. P. Eckstein. Visual search: A retrospective. *Journal of Vision*, 11(5):14–14, Dec. 2011. doi: 10.1167/11.5.14 2

[6] U. Gruenefeld, D. Ennenga, A. E. Ali, W. Heuten, and S. Boll. EyeSee360: designing a visualization technique for out-of-view objects in head-mounted augmented reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, pp. 109–118. ACM, Brighton United Kingdom, Oct. 2017. doi: 10.1145/3131277.3132175 2

[7] X. Hu, A. Moore, J. Coleman Eubanks, A. Aiyaz, and R. P. McMahan. Evaluating Interaction Cue Purpose and Timing for Learning and Retaining Virtual Reality Training. In *Symposium on Spatial User Interaction*, pp. 1–9. ACM, Virtual Event Canada, Oct. 2020. doi: 10.1145/3385959.3418448 1

[8] X. Hu, A. G. Moore, J. C. Eubanks, A. A. Aiyaz, and R. P. McMahan. The Effects of Delayed Interaction Cues in Virtual Reality Training. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 63–69. IEEE, Atlanta, GA, USA, Mar. 2020. doi: 10.1109/VRW50115.2020.00019 1

[9] B. Kelley, C. Wickens, B. Clegg, F. R. Ortega, and A. C. Warden. Guiding Gaze: Comparing Cues for Visual Search. In *The 31st IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE, Orlando, FL, USA, Mar. 2024. doi: 10.1109/VRW62533.2024.00323 1, 2

[10] B. Kelley, C. Wickens, A. C. Warden, B. Clegg, and F. Ortega. 2D and 3D Augmented Reality Attention Cueing Comparisons in 3D Target Search. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, p. 10711813241265077, Aug. 2024. doi: 10.1177/10711813241265077 1, 2

[11] Y.-J. Kim, R. Kumaran, E. Sayyad, A. Milner, T. Bullock, B. Giesbrecht, and T. Hollerer. Investigating Search Among Physical and Virtual Objects Under Different Lighting Conditions. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3788–3798, Nov. 2022. doi: 10.1109/TVCG.2022.3203093 2

[12] R. Kumaran, Y.-J. Kim, A. E. Milner, T. Bullock, B. Giesbrecht, and T. Höllerer. The Impact of Navigation Aids on Search Performance and Object Recall in Wide-Area Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–17. ACM, Hamburg Germany, Apr. 2023. doi: 10.1145/3544548.3581413 2

[13] D. Mifsud, C. Wickens, M. Maulbeck, P. Crane, and F. R. Ortega. The Effectiveness of Gaze Guidance Lines in supporting JTAC's Attention Allocation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1):2198–2201, Sept. 2022. doi: 10.1177/1071181322661143 1, 2

[14] N. Moacdieh and N. Sarter. Display Clutter: A Review of Definitions and Measurement Techniques. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(1):61–100, Feb. 2015. doi: 10.1177/0018720814541145 1, 3

[15] U. Neisser. VISUAL SEARCH. *Scientific American*, 210(6):94–103, 1964. Publisher: Scientific American, a division of Nature America, Inc. 2

[16] H. Oshimi, M. Perusquía-Hernández, N. Isoyama, H. Uchiyama, and K. Kiyokawa. LocatAR: An AR Object Search Assistance System for a Shared Space. In *Augmented Humans Conference*, pp. 66–76. ACM, Glasgow United Kingdom, Mar. 2023. doi: 10.1145/3582700.3582712 2

[17] A. Raikwar, D. Mifsud, C. D. Wickens, A. U. Batmaz, A. C. Warden, B. Kelley, B. A. Clegg, and F. R. Ortega. 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. *IEEE Transactions on Visualization and Computer Graphics*, 30(5):2662–2670, May 2024. doi: 10.1109/TVCG.2024.3372062 1, 2, 3

[18] A. Raikwar, D. Mifsud, C. D. Wickens, B. Kelly, A. C. Warden, B. A. Clegg, and F. R. Ortega. Beyond the Wizard of Oz: Using Imperfect Machine Learning to Examine the Impact of Reliability of Augmented Reality Cues on Visual Search Performance. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 437–438. IEEE, Sydney, Australia, Oct. 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00093 1, 2

[19] A. Seeliger, R. P. Weibel, and S. Feuerriegel. Context-Adaptive Visual Cues for Safe Navigation in Augmented Reality Using Machine Learning. *International Journal of Human–Computer Interaction*, 40(3):761–781, Feb. 2024. doi: 10.1080/10447318.2022.2122114 3

[20] L. J. Skitka, K. L. Mosier, and M. Burdick. Does automation bias decision-making? *International Journal of Human-Computer Studies*, 51(5):991–1006, Nov. 1999. doi: 10.1006/ijhc.1999.0252 1, 2

[21] L. J. Skitka, K. L. Mosier, M. Burdick, and B. Rosenblatt. Automation Bias and Errors: Are Crews Better Than Individuals? *The International Journal of Aviation Psychology*, 10(1):85–97, Jan. 2000. doi: 10.1207/S15327108IJAP1001_5 1, 2

[22] J. Sweller. Cognitive Load Theory. In *Psychology of Learning and Motivation*, vol. 55, pp. 37–76. Elsevier, 2011. doi: 10.1016/B978-0-12-387691-1.00002-8 1

[23] A. C. Warden, C. D. Wickens, D. Rehberg, F. R. Ortega, and B. A. Clegg. 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, Dec. 2023. doi: 10.1109/THMS.2023.3302152 1, 2, 3

[24] Weiquan Lu, H. B.-L. Duh, S. Feiner, and Qi Zhao. Attributes of Subtle Cues for Facilitating Visual Search in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 20(3):404–412, Mar. 2014. doi: 10.1109/TVCG.2013.241 2

[25] C. D. Wickens, R. S. Gutzwiller, and J. S. McCarley. *Applied attention theory*. CRC Press, Taylor & Francis Group, Boca Raton, second edition ed., 2023. 1, 3

[26] T. L. Y. Wu, A. Murphy, C. Chen, and D. Kulić. Adaptive cueing strategy for gait modification: A case study using auditory cues. *Frontiers in Neurorobotics*, 17:1127033, Mar. 2023. doi: 10.3389/fnbot.2023.1127033 3

[27] Y. Zhao, E. Kupferstein, B. V. Castro, S. Feiner, and S. Azenkot. Designing AR Visualizations to Facilitate Stair Navigation for People with Low Vision. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 387–402. ACM, New Orleans LA USA, Oct. 2019. doi: 10.1145/3332165.3347906 2