THESIS

GUIDING GAZE, EVALUATING VISUAL CUE DESIGNS FOR AUGMENTED REALITY

Submitted by

Brendan Kelley

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2024

Master's Committee:

Advisor: Francisco R. Ortega

Cyane Tornatzky
Dr. Mohammed Safayet Arefin

Copyright by Brendan Kelley 2024

All Rights Reserved

ABSTRACT

GUIDING GAZE, EVALUATING VISUAL CUE DESIGNS FOR AUGMENTED REALITY

Visual cueing is an interdisciplinary and complex topic. It has garnered interest for implementation with extended reality (XR). Both augmented reality (AR) and virtual reality (VR), are often employed for visual search tasks. Visual search, a paradigm rooted in cognitive psychology (in particular attention theory), can often benefit from cueing interventions. However, there are several potential pitfalls with using cueing techniques in AR; namely, automation bias, clutter, and cognitive overload. These factors are tied to design and implementation choices, such as modality, representation, dimensionality, reference frame, conveyed information, purpose, markedness, or the task domain. Design factors are subject to both the cognitive factors, as well as, technical specifications of the display technology. To address these factors, this work proposes a within-subject four factor design addressing the question *how do different cue designs affect visual search performance?* Four cueing conditions are used: no cue (baseline), gaze line, 2D wedge, and 3D arrow. Results support the use of cues for visual search, however the gaze line condition provided for the fastest search time, accuracy, and greatest reduction in head rotation. Additionally, the gaze line cue was preferred by participants and was produced more favorable NASA TLX scores.

ACKNOWLEDGEMENTS

Thanks to Dr. Ortega and the NUI Lab for the constant support and assistance. Thanks to Cyane Tornatzky for always being a fantastic collaborator. This study was partially supported by ONR grants N00014-24-1-2214 NSF 2238313, N00012-21-1-2949, and N00014-21-1-2580.

DEDICATION

For my cat, Chewy, who has always been there for me. For my partner, Alissa, for all of the love and support during my program. For my parents, for supporting me through all of these years. For my brother, who always has my back. For my uncle, who's reckless donation of funds towards my first gaming console, set me on this path. Für meine Oma. Ich hoffe, dass sie stolz auf mich ist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

As extended reality (XR) technologies have continued to advance and develop we have seen their use applied to a variety of tasks. Training and education [9], assembly directions [10], and visual search [11] tasks have all seen successful adoption of both augmented reality (AR) and virtual reality (VR) technologies. To assist in tasks, visual cueing techniques are often employed such as 2D arrows [12], spatial highlights [13], or other novel approaches specific to a problem [14] or task [15]. In particular, visual cueing techniques have been shown to increase overall performance with visual search.

However, not all cues are created equal. A variety of different design factors affect cue implementation including: modality, representation, dimensionality, the reference frame, conveyed information, cue purpose, markedness, and contextual domain [16–19]. Each of these factors affects the overall benefit and efficacy of the cueing intervention; with the goal often being to decrease search time and increase accuracy. In addition to these design factors, cognitive aspects must also be considered for visual cueing. Each of these factors is considered to explore the question: *how do different cue designs affect visual search performance?*

Attention theory [20, 21] can help to explain the underlying cognitive factors that influence performance with a given cue. In particular, aspects such as automation bias, clutter, and cognitive overload are crucial to consider. These factors can quickly erode the potential benefits of visual cueing techniques. As such both the cognitive basis for visual search, as well as, the design factors influencing XR cueing must be considered.

To explore visual cueing for AR, a testbed application was developed using Unity (version 2022.2.1f [22]) for the Magic Leap 2 AR head mounted display (HMD). Three cues were utilized (gaze line, 3D arrow, and 2D wedge) along with a baseline no cue condition. Each cue was specifically chosen based on pilot study results, as well as, their performance or prevalence in prior literature, and how the communicate spatial information. Fourteen participants used each of these cue conditions to search for a target among a series of distractors in a virtual search environment projected over an ArUco AR marker. Results from this study addressed the question, *how do different cue designs affect visual search performance?* Additionally, results demonstrate a benefit from cueing, with the gaze line cue providing for the fastest response time, highest accuracy, least amount of head rotation, and overall user preference.

Chapter 2

Literature Review

2.1 Augmented Reality for Visual Search

Visual search is a common, everyday task that individuals engage with. For instance, looking around a crowded bar for a friend, looking around a room for car keys, or scanning for keywords in a thesis document. Each of these is an example of visual search; wherein an individual is scanning an environment, be that a bar, room, or text, for a target [23]. Often there are other distractors that hinder the search. These distractors may be similar objects, such as people, or other salient features of the environment [24].

The use of XR for visual search tasks has proven to yield beneficial results, with several studies finding that the use of cues improved search time and accuracy metrics [25–27]. A variety of other factors have also been considered in relation to both the affordances and limitations of current AR HMDs including lighting [28], out-of-view targets [14, 29, 30], multimodal approaches [31, 32], collaborative cueing (especially with expert oversight) [33,34], and more. However, these technical design factors are still rooted in cognitive theory creating a unique, often muddled, interdisciplinary problem space to explore.

2.1.1 Key Terms

Due to the interdisciplinary nature of cueing research several terms often become conflated. For instance, the terms point-of-interest [35, 36], or area-of-interest [37] can often be used in lieu of the term target. Similarly, cues can also be referenced as visualizations [38], or aids [11]. For the purposes of this discussion the terms target and cue are defined as follows:

- **Target:** The object, place, thing, or stimuli being searched for. Often used interchangeably with area-of-interest, or point-of-interest.

Figure 2.1: An example of a classic visual search task, such as those inspired by [1,2]. In this the participant would be asked to search for the target k among the distractor x's.

- **Cue:** Output from the XR system designed to facilitate a response, incite behaviour, or assist in a task. Often used interchangeably with aid or visualization.

While the interdisciplinary nature of XR cueing for visual search does provide for a rich problem area to explore, the conflated nature of terms and definitions can lead to confusion and challenges between each of the fields. Thus the importance of explicating these terms.

2.1.2 Visual Search

The visual search paradigm in cognitive psychology focuses around identifying targets within a search field, potentially with the presence of distractors. This process is rooted in information processing [1], human cognition, and attention [21, 39]. In classical visual search tasks, such as the example in Figure 2.1, participants are shown a slide or projection with a series of shapes, characters, or figures. They are then asked to find a specific target from among these [1, 2]. In the case of Figure 2.1, a participant would be asked to find the green "k" from among the distractors.

Visual search can be affected by a wide variety of factors, including the search stimuli, the user's methods for search, environmental guidance (such as cues), perceived value, and prior search history [40]. With "bottom-up" guidance, aspects such as color, irregular-shape, or other aspects of the search stimuli draw attention [41]. For instance if searching for a specific playing card (i.e. Jack of Hearts) on a table the color and shape of the symbols would serve to guide attention for searching; as the black symbols (club and spade) and the pointed red diamonds can be ignored. A "top-down" approach for attention guidance in contrast focuses on objects that closely resemble the target stimuli, with little to no attention directed at distractors that are of a greatly divergent design [42]. It is also important to note that "bottom-up" and "top-down" control for visual search can be combined in a task, and that the switch between the two can be an automatic, unconscious process [43].

In addition to "bottom-up" and "top-down" guidance, the perceived value of the search items, the individual's prior experience with search, and aspects of the scene (or environment) can all affect the search process [40]. By assigning a perceived monetary value to the selection of otherwise arbitrary, neutral, distractors Anderson et al. [44] were able to manipulate a search task by guiding focus away from the desired search stimuli. This phenomenon is also in part learned. As participants were exposed to the perceived monetary value of certain selections, they began to seek out those selections over the target stimuli (which also was assigned monetary value) [44]. Indeed memory becomes a useful factor for understanding search.

Both semantic and episodic memory can inform an individual's search process [45]. Semantic memory provides guidance for search based on logical assumptions on where an object typically resides in a scene. For instance, a cup sitting on a table rather than a sofa. Episodic memory provides guidance based on the last time a particular scene was encountered. For instance, recalling placing keys on a desk. These factors also relate to the notion of scene guidance for visual search [46]. Even just a brief glimpse of a scene can create mental constructs that are used to guide search based on where an individual would assume items to be (i.e. cup on table, pen on desk, etc.) [47].

2.2 Theoretical Elements

"Bottom-up," "top-down," history (or priming), reward, and scene context are all based in human cognition [48]. Attention theory [20,21,39] provides useful context for human performance during any given task; AR or otherwise. Both theories also explicate on potential limitations, shortcomings, and dangers with AR cue integration; namely cognitive overload, clutter effects, and automation bias.

2.2.1 Cognitive Theories

Attention theory has key foundational assumptions that inform and explicate their more robust concepts. In 1968 Atkinson and Shiffrin proposed the Multi-Store Memory Model (see Fig. 2.2) [3]. This model proposes that environmental stimuli is first passed to sensory memory. In sensory memory the environmental information can either be attended to (via mechanisms proposed by attention theory [20, 21, 39]) or forgotten due to a variety of factors, such as cognitive overload, irrelevance, attention filtering, etc. From here, any information attended to would be passed to short term memory where that information is rehearsed. If the information is useful, it may then be passed to long term memory for later retrieval when exposed to new experiences. Information in the short term memory may also be forgotten if not rehearsed, or if it is deemed as unnecessary for the task at hand. Similarly, even when committed to long term memory, information may fade over time when not recalled. This Multi-Store Model informs a variety of other cognitive models, including the Limited Capacity Model of Motivated Mediated Message Processing (or LC4MP) [49, 50], attention theory [20, 21, 39], cognitive load theory [51, 52], and the Working Memory Model [4].

The Working Memory Model [4] builds upon the Multi-Store Model [3] by introducing three key concepts: the Central Executive, the Visuospatial Sketchpad, and the Phonological Loop (see Fig.2.3). The key function of the Central Executive component is to delegate information and resources between long term memory (LTM), the Visuospatial Sketchpad, and the Phonological Store [53]. The Central Executive also helps to focus attention, and even divide attention, among

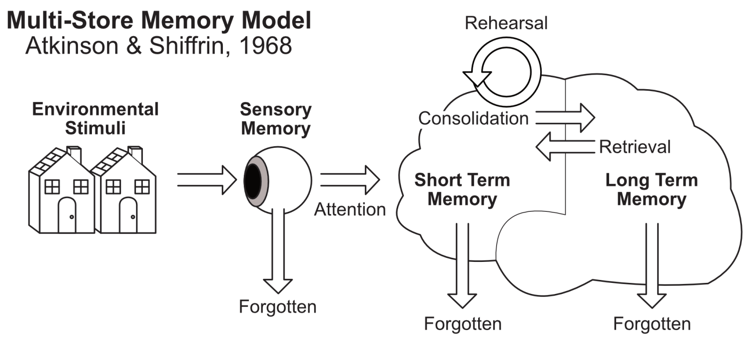

Figure 2.2: The Multi-Store Model of memory as proposed by Atkinson & Shiffrin [3].

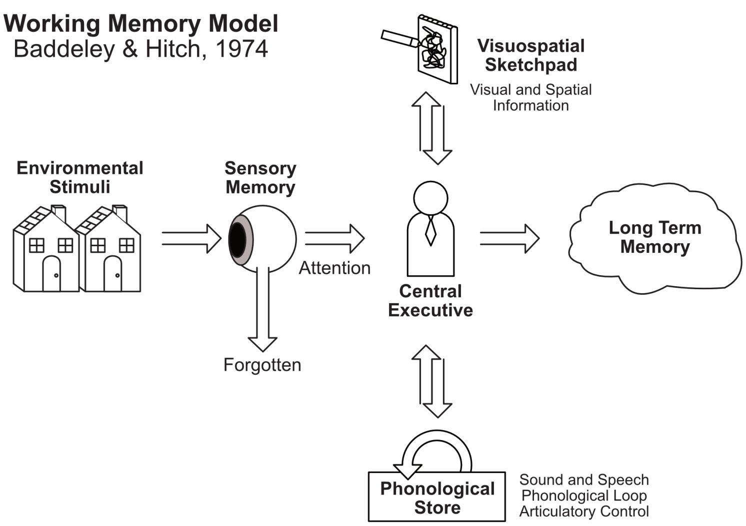

Figure 2.3: The Working Memory Model as proposed by Baddeley & Hitch, 1974 [4].

multiple tasks [53]. The Visuospatial Sketchpad is used to process visual and spatial information [4, 54]. The Phonological Store, on the other hand, processes speech [4, 54, 55]. Each of these functions works in tandem with the Central Executive to dictate what is stored in LTM and what is forgotten. As with the Multi-Store Model information can also fade from LTM over time. Both of these models (Multi-Store Model and Working Memory Model) draw from and inform attention theory. This model is also supported by neuroscience works over the years [56].

Attention theory is actually a collection of different theories that pertain to human attention [20, 21]. It describes a wide range of different mechanics that drive attention, including attention as both a filter and a resource [39] and the link between perception and memory [57].

As a filter attention serves as a selector, that controls the flow of information from the environment for processing. When used as a filter attention helps to prioritize pertinent information amongst competing distractors and information channels [33]. This process is typically referred to as selective attention. This process helps to drive what spatial and environmental information is exposed to working memory, however just because something is perceived to does not mean it will also be attended to [58]. Humans actually neglect to commit a vast amount of information experienced daily to memory. This filtering function of attention helps to dictate not only what reaches working memory, but what is also committed to memory, and is a potential mechanism in the sensory memory component of the Multi-Store [3] and Working Memory models [4].

Working memory is conceived as a small amount of cognitive resources that store information in a ready to use form [59]. Information stored in working memory is then used to approach an immediate task, problem, or activity and is not stored for an extended period of time [60, 61]. Instead information in working memory is either discarded or committed to long term memory for later retrieval [62]. Several other theories help to explain the mechanism behind long term memory; for instance schema theory [63, 64]. This theory posits that individuals use schema to build hierarchical relationships with other schema. This produces a web of understanding with each schema being referential to multiple other schema. With each new interaction we may reinforce or build new schema [65]. Schema theory is also closely tied to semantics, where a signifier relates to and conjures information on whatever is signified (i.e. the word tree relates to a more specific schema pertaining to pine trees) [66].

Attention can also be conceptualized as a resource [39]. When considered as a resource, attention serves as a mediating factor to determine what is committed to memory. In both of these cases, attention or cognitive resources are limited. The exact allocation and availability of these resources is both task dependent and mediated by individual factors [62,67]. For instance, if a search task only required finding a single target, a large amount of cognitive resources could be dedicated solely to finding that target, however if a second target was present that would divert some resource to that second target. Additionally, if a secondary task were introduced, such as wayfinding, that task would then further divert resources to the completion of the secondary task. If an individuals overall workload increases beyond the capacity of cognitive resources currently available, performance will begin to suffer [68]. This phenomenon, referred to as cognitive overload, is a common issue when using technological interventions such as AR cueing.

2.2.2 Cognitive Limitations

Cognitive overload becomes a key danger when considering the completion of any task. Because individuals have a limited amount of cognitive resources, any task that demands more resources than are currently available will lead to overload [52]. Similarly, if an individual is not sufficiently stimulated performance may still suffer [51]. This leads to the notion of the *Goldilocks zone*, where just the right amount of cognitive resources must be allocated to achieve peak performance [20, 39]. AR cueing can aid in reducing cognitive load by reducing the cognitive resources require to attend to and search for a target, however, other factors such as prior knowledge [10, 69] may also impact task performance.

While AR cueing can be beneficial, cues may also induce clutter effects, which can negatively impact performance. Clutter can be a product of both the cue design (screen clutter) or a product of the environment (environmental clutter) when considering AR. In the case of screen clutter, this is a product of the virtual elements (often situated on a 2D UI plane). If the amount of items, size of items, or salience of items presented virtually to the user requires too much attention to process this may lead to clutter effects. Similarly, if the environment is exceptionally busy (i.e. bright packaging in a supermarket), this may conflict with the cue designs. For instance, if the cue is a green arrow and the search environment is a green forest the cue may be difficult to discern from the environment.

These clutter effects emerge when the information (i.e. cues) presented in the AR HMD begin to require extra cognitive resources to process [70, 71]. By introducing too many cues within a limited space (such as a low field of view, FOV, AR HMD) users have to take time and utilize resources to discern and differentiate between different cues, as well as their environment, the targets, and distractors. Cues can be designed to minimize these effects, such as with the wedge cue [72], which was originally designed to minimize clutter on GPS screens. However, they can also contribute to these effects when cues become too information dense [25, 73].

Clutter and overload effects can also go hand in hand with automation bias. Automation bias emerges when users rely too heavily on the automated system (such as AR cues). When utilizing perfect cues (where the cue always points to the target without error) the cue with the highest performance will often suffer from automation bias once errors are introduced [74–76]. This is often due to the saliency of the cue. A cue's salience is predicated on a variety of factors (and can be mediated by environmental factors such as environment clutter). Color, motion, size, and a host of other design factors can contribute to salience. When a cue has a high level of saliency it will draw attention away from the environment, targets, and task.

With high cue salience it becomes far to appealing and easy to blindly follow the cue recommendations. These effects can even persist when working in groups [77]. When automation bias effects emerge users trust the system over their own perceptions leading them to follow the automated recommendation even when that recommendation is incorrect. This in turn, increases the number of errors during the completion of a task [78, 79]. This can prove especially dangerous in high intensity or hazardous environments, such as that of an airplane cockpit [76, 79]. Automation bias may also be the produce of humans' tendency toward being "cognitive misers." This is the notion that individuals try to expend as little cognitive resources to complete a task as possible [80]. If a cue offloads too much cognitive processing, it may become far more enticing to blindly trust the cued decision over attending to the environment and task at hand.

Attention theory, clutter, and automation bias become important aspects when designing and implementing AR cueing interventions for visual search as they underpin a variety of different human factors that will inevitably intertwine with the cue design. As such the concepts explicated by these theories should be considered when working with and design systems for AR cueing.

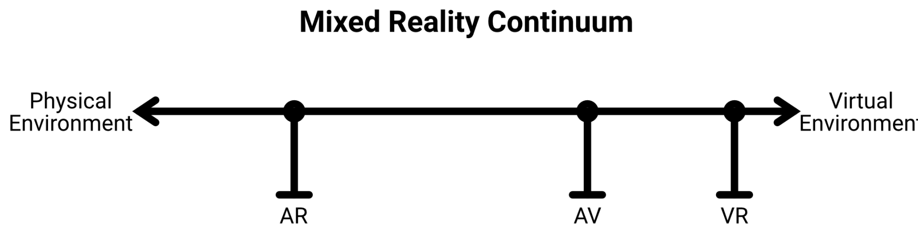

Figure 2.4: The mixed reality continuum as proposed by Milgram & Kishino [5].

2.3 Elements of Cue Design

Several important elements of cue designs have been identified through prior exploration [16–18, 81, 82]. These elements include *modality*, *representation*, *dimensionality*, *reference frame*, *conveyed information*, *purpose*, and *markedness*. Additional context surrounding the cue implementation is also important, as different cueing techniques are often developed with a specific task in mind (i.e. wayfinding [83, 84] vs. visual search [25, 85]). This takes the form of a cue's *domain*.

2.3.1 AR vs VR Cueing

This discussion, and subsequent study, are predominately focused around AR, however VR has also been used extensively for cueing. Both AR and VR are situated along the mixed reality continuum (see Fig.2.4) [5]. AR aims to blend virtual objects into the physical environment [86], whereas VR aims to suppress the physical in favor of the virtual [87]. In spite of these difference, there is evidence to suggest that simulated AR in VR can yield comparable results to using a live AR system. Lee et al. [88] replicated the work of Ellis et al. [89] where participants guided a virtual ring through a virtual path. The original study by Ellis et al. [89] utilized AR, whereas the replicated work by Lee et al. [88] utilized VR and produced comparable results to the original study.

Simulated AR was further validated in another study by Lee et al. [90], where participants traced a yellow pipe in a room with 3 orange pipe distractors. Each pipe intersected with various letter targets. As participants traced the pipe they verbally indicated what letter was encountered. Both AR and VR were used with 3 different tracking latencies (0ms, 50ms, and 150ms). Results demonstrated, that so long as latency was within one standard deviation of the AR condition, VR again produced comparable results. As such, results and experiments from VR can inform AR design and *vice-versa*. Several other studies have used VR to test AR cue designs including visual search [37, 91] and navigation tasks [92].

2.3.2 Modality

Modality refers to the type of sensory stimuli used for cueing. This may be visual, auditory, tactile, forceful, olfactory, or gustatory in nature. Cues may also be multi-modal; adopting multiple different stimuli for cueing.

Visual cues are by far the most common. This is due to the dominance of vision in everyday lives; approximately 80% of all environmental information is processed visually [93, 94]. Additionally, most AR and VR HMDs are build with two main forms of display feedback: visual and auditory. This in turn, makes those two modalities (audio and visual) much easier to implement. Visual cues do sufferer from a variety of environmental factors. Lighting, background cutter/salience, and other visual aspects can hinder visual cueing techniques [28,95]. An additional legacy bias also exists as many cue designs are take from 2D display counterparts which primarily afforded visual techniques. Halo, for instance, was originally designed for car GPS devices [96], as was the wedge cueing technique [72].

Auditory cues are also a product of HMD designs as most HMDs contain speakers or headphones. With auditory cues, sound becomes the primary modality for conveying information. However, auditory cues are often used in conjunction with other modalities such as visual [27, 32, 84]. Often these are paired with visual techniques, but examples of other multi-modal approaches exist such as audio-tactile [27]. Often audio is used to denote transitions or events rather than constant information, like a target coming into view [73] or entering a hazard area [32].

Tactile, force, olfactory, and gustatory are not nearly as prevalent as visual and auditory cues, but do provide opportunities for unique and accessible interactions. Both tactile and force cues depend on the sensation of touch, however tactile is often conceptualized as surface touch, whereas force involves pressure. For instance the stylo and handifact design [6] utilize force by adjusting

Figure 2.5: The stylo device and the handifact cue in action. Taken from [6].

the strength of feedback from the custom stylo device. This is paired with a virtual hand that appears to "push" on the user to try and guide them to the proper body position (see Fig. 2.5). For this design both tactile and force feedback is important, whereas another study on simulating roughness only utilized tactile cues for their design [97]. Different tactile sensations were paired with virtual representations to convey surface texture in VR. While 10 different materials were explored, only two levels of simulation were needed to simulate a variety of textures [97].

As with tactile and force cues, gustatory and olfactory cues often require a custom ancillary device for cueing; as HMDs often cannot produce tastes or smells. These modalities are often paired together, as our sense of taste is tied to our sense of smell [98, 99]. While the use of olfactory and gustatory cues is possible they are often overlooked, in part due to struggles identifying opportunities for implementation [100].

2.3.3 Representation

Representation is how information is presented to the user. This may be alphanumeric, symbolic, captured, characterized, via replica, or emphatic. Traditionally these categories have been primarily been explored in the visual modality [101], however these categories can also apply to non visual modalities (such as speech as alphanumeric or sound recordings as captured).

Alphanumeric cues use characters, words, or phrases to present information. This may be visual text, such as number of meters from a target, or recorded speech directions [84]. Symbolic representation uses some form of symbol to indicate information. A common example of this is the use of arrow cues, where the arrow serves as a symbol to indicate direction towards the target [25, 102, 103]. Captured, characterized, and replicas, all draw from the real world target in some way, albeit with various levels of abstraction. Captured uses photos, videos, or sound clips of targets. Characterized cues are abstracted one step further, by presenting the target using drawings, diagrams, blueprints, etc., such as with Laviola et al. [104] where a technical diagram of a part was displayed to cue a maintenance task. Replica is similar to captured, but uses models or animations of the target object instead of images or recordings, much like a digital twin [105]. Emphatic cues affect, add to, or alter the physical target to indicate some aspect of it (most often position), such as spatial highlights [26].

2.3.4 Dimensionality

Dimensionality defines the number of dimensions used to convey information. As AR HMDs are 3D displays, and typically part of a 3D user interface [81, 87], the dimensionality of a cue is composed of 1 - 3 total dimensions. 2D and 3D cues are the most commonly implemented cues, however 1D cues are still possible such as a fuel gauge used to represent visibility of a target object from the user's position for an architectural planning application [106]. Here the only metric is visibility (ranging from 0 to 100 percent visible) and the gauge fills along one axis or dimension.

Cues using 2 dimensions are typically placed on a user interface (UI) plane. These translate some aspect of the cue into 2D representation. By far the most common example is that of the 2D arrow which is place on the UI and rotates to point to the target from the users current view. Three dimensional cues, such as the 3D arrow, 3D wedge, and 3D radar used in Yu et al. [91] are 3D objects that move in a local 3D space to represent some aspect of the target and user.

2.3.5 Reference Frame

The reference frame of a cue refers to the nature of spatial information used to drive that cue. This may be egocentric (in reference to the self) or exocentric (in reference to the external environment) [107]. Egocentric category cues may be further organized based on the specific aspect of the user they reference, such as station point, retinocentric, headcentric, bodycentric, or proprioceptive [19, 81, 107].

With most egocentric reference frames the station point, retinocentric, headcentric, or bodycentric frames are used. This is due to them referencing the eyes (station point and retinocentric), head (headcentric), or entirety (bodycentric) of the user, all of which are easily accessed when developing for AR HMD systems. Vision referenced and head referenced cueing are especially useful for collaborative experiences as these designs allow users to explicitly see what the other person is looking at [33, 108]. Proprioceptive is not as widely used, as it typically depends on tactile or force modalities which often require custom feedback devices such as the stylo [6] or VR walking canes [109, 110].

Exocentric references typically center on the target. This may be through highlights, arrows, or other emphasized cues. For instance a study on locating buttons in a physical environment overlaid cubes or squares onto the physical buttons [26]. In this case the cube is centered around the exocentric reference of the search target. Often cues draw from both exocentric and egocentric reference frames such as with the gaze line design [111].

2.3.6 Conveyed Information

The direction to, position of, proximity to, transition, or other context specific information can all be conveyed by a cue. Cues may also convey layers of information. For instance, the 2D arrow conveys just the direction of a target [96], whereas a 2D wedge conveys direction, distance, and position [72, 112].

Direction and position are the most commonly communicated bits of information when using cues. These are often used for visual search tasks where users must find a target in an environment (with or without distractors) [20, 21, 23–25, 74]. This could be finding office supplies in shared spaces [113], searching for targets from above (such as in a plane) [114], or identifying areas of interest in a room [11, 115]. Many tasks also include visual search, even if simply finding the target is not the end goal. For instance a manipulation task, such as moving multiple objects in an unordered fashion, requires that users first identify the target to be manipulated [15].

2.3.7 Purpose

The purpose of the cue relates to the task being completed. In particular, the action or state that causes the cue to disappear, or the exit condition. Discover, look, go, select, manipulate, operate, gesture, and speak are common purposes for cues [16, 17, 82]. These relate to the task being completed, but do not necessarily encompass the entirety of the task. Some tasks, such as select, may first require the user to discover and look before the final selection. In cases such as these selection is the purpose, but other aspects are also present to complete that goal.

Discover cues help the user to search for and find a target. These are often used for visual search tasks [21, 23, 24, 116]. This could be office supplies [113] or ground targets [114], in any case the goal is to find and discover the target. Look cues are build upon discover cues. After discovering a target, the user needs to view or inspect the target. For instance, in a wide-area search, participants needed to look at the target gemstones to make judgements regarding their surface textures [11].

Go cues require the user to relocate to the target location. This is unlike discover and go cues, which typically don't require the user to relocate. Wayfinding and navigation tasks often utilize go cues. Both map and arrow cues have been shown to improve wayfinding and navigation performance [83]. These cases often focus on wayfinding even if a secondary task is also in progress [117].

Select, manipulate, and operate cues all require the user to affect or use the target in some way. Selection is the most common, especially in relation to visual search. For visual search selection tasks, once a target has been found the user must select it [30, 108, 114]. Manipulate tasks, on the other hand, require the user to manipulate the target in some way after selection.

Most often this requires rotating, positioning, or scaling the target. For instance, in a bimanual unordered task users had to manipulate target objects positions after selecting [15], or in a remote collaborative assembly task, one user positioned pieces to the correct position and rotated them to the proper orientation [34]. Manipulation first requires discovery and selection of a target, but the final operation is to reposition, rotate, or scale the selected target. Operate is similar to manipulate, but depends on specific functions or elements of the target (i.e. turning on a flashlight using the power button). For these, the cue prompts user to use a specific property or function of the target, such as pressing a button [26] or removing bolts [104].

Gesture and speech cues aim to solicit responses from the user. Gesture cues ask the user to perform a movement, pose, or action. This could be for martial arts or other physical training [6]. This could also be for physical interventions, such as using cues for gait control for individuals with Parkinson's [118]. Speech cues prompt vocalizations from the user. This could be words or phrases, but may also include general noises. These are often used for secondary tasks, such as measuring cognitive load [11].

2.3.8 Markedness

Markedness is the extent to which the cue blends into the environment. Cues may be subtle, emphasized, integrated, or overlaid. Subtle cues blend into the environment. With these, there is little differentiation between the cue and the environment, or target. These are not commonly used, but do exist. Often, subtle cues take advantage of our physical world and perception; such as using shadows or parallax effects for depth in surgical operations [119]. Others are simply manipulations in lighting or color of the target [115].

Emphasized and integrated cues are much more readily apparent. These change the environment in an overt manner, but emphasized cues change an aspect of the physical object, whereas integrated cues add a new virtual object to the environment. Emphasized cues include highlighting a pedestrian crossing a street with a yellow highlight [120], changing the color of the target [121, 122], or placing colored wireframes over a button [26]. Integrated cues include attention tunnels/funnels [123–125] and bezier arrows [85, 126].

Overlaid cues are unique in that they are almost always 2D, and exist on the 2D UI plane of a HMD. These may extrapolate from 3D information, but are always 2D or 1D in nature. These include minimaps [25, 108], 2D compasses [11], and 2D radar designs [11].

2.3.9 Domain

Domain relates to the specific context surrounding a cue's use. Some cues are developed with a specific task or purpose in mind, which dramatically shapes their implementation. For instance the Search and Rescue AR Visualization Environment (SAVE) implements a series of visual cues, such as frame highlighting and drawing tools, for first responder use [127]. Another tool, MRsive, focuses on wayfinding and engagement for art museums [128]. Another study built a variety of cues specifically for robot teleoperations [129].

Each of these design practices, along with the cognitive aspects of visual search and cueing, help to inform the presented study. These aspects have been carefully considered while developing the testbed application and experiment focused on exploring the question *how do different cue designs affect visual search performance?* As discussed in Section 2.3 a variety of technical implementation aspects can dramatically affect cued tasks. These can be further influenced by the cognitive aspects related to attention theory.

2.3.10 Concluding Summary

Visual cueing for AR benefits from a rich background. Cognitive theories [20, 21, 51, 52], technical design [16–18, 81, 82], mediated theories [50], and more all help to inform designs, implementation, and use cases for visual cueing. This can lead to conflating terms, such as target with point-of-interest or area-of-interest, or cue with visualization or aid.

Much of visual cueing research has been applied to tasks that include visual search [1, 2, 23, 24, 40, 41]. This paradigm is rooted in cognitive psychology, and typically asks an individual to find a target in an environment among distractors. Both the Multi-Store Memory Model [3] and the Working Memory Model [53] help to explain the mechanisms underpinning attention theory [20, 21, 39]. Each of these highlights the important human elements that affect the efficacy and performance of search. Especially, the dangers posed by automation bias [76, 77], cognitive overload [51, 52], and cluttering effects [70, 71].

These are invariable tied to the specifics of cue design. Modality, representation, dimensionality, reference frame, conveyed information, purpose, markedness, and domain can all affect the efficacy and usefulness of a cue [16–18, 81, 82]. Efficient cue designs tie these factors into the cognitive aspects of cueing, but are also limited by the technical elements of the display technology; in this case XR HMDs. Each of these factors suggests the research question *how do different cue designs affect visual search performance?*

Research Question and Hypothesis

In addition to the research question: *how do different cue designs affect visual search performance?* Prior work suggests several hypothesis that merit exploration.

The cognitive factors suggested by attention theory [20, 21] suggest two hypotheses. Clutter becomes an important factor to consider, as cues that take up more of the users FOV may suffer from cluttering effects [70, 71]. This suggests H2 *Cues that take up more of the user's FOV will be less effective.* Salience becomes another important factor, as cues that are demand more attention from the user may become very compelling. With perfectly cued targets this should result in beneficial outcomes, but may lead to detrimental effects when errors are introduced [76, 77]. Even with error, but especially with perfect cueing, any cue should improve upon the baseline no cue condition. This suggests H3 *The more salient the cue design the more effective it will be.*

- H1: Any cue condition will improve upon the baseline condition.
- H2: Cues that take up more of the user's FOV will be less effective.
- H3: The more salient the cue design the more effective it will be.

The design of cues also suggests several hypotheses [16–18, 81, 82]. In particular, the embedded information can be a beneficial when it reduces the amount of processing an individual must do to locate targets (i.e. a cue that only shows target direction would require cognitive resources to extrapolate location). As such cues with more embedded information should provide for better results. This suggests H4 through H7:

- H4: Cues with more embedded information will produce faster search times.
- H5: Cues with more embedded information will produce higher accuracy rates.
- H6: Cues with more embedded information will reduce user movement.
- H7: Cues with more embedded information will reduce task demand.

To address these hypotheses, time, accuracy, and head movement data are collected. Additional NASA TLX scores and rankings are collected to address H7 specifically.

Chapter 3

Methods

To address the question *how do different cue designs affect visual search performance?* a comparative research study is used. A testbed application was developed for use with the Magic Leap 2 AR HMD using Unity (version 2022.2.1f [22]). Two pilot studies were conducted to evaluate the cue designs as well as the experiment design. Pilot study 1 focused on the 8 original cue designs implemented. Performance data and feedback from participants was then used to modify the application for pilot study 2 and narrow the selection of cues (down to 3). Pilot study 2 focused on the experiment design, testing the full procedure used in this study. Changes from these two pilot studies helped to validate the testbed application as well as verify the experiment design and procedure.

3.1 Apparatus

The testbed application used for this study was developed using Unity. It is compatible with most AR HMDs, but for this study the Magic Leap 2 was utilized. The Magic Leap 2 HMD also includes one motion controller that was used for selecting targets during the experiment using the trigger button and a ray interactor. A standalone *.apk* file was loaded onto the HMD and the experiment ran solely on the Magic Leap 2 hardware. The virtual elements were spawned using ArCuo markers. To configure and drive the visual search task a configuration text document was loaded that contained 64 total search configurations.

3.1.1 Cue Designs

While eight different cues were implemented into the test bed application only 3 of them were selected for this experiment: gaze line, 2D wedge, and 3D arrow. Each was selected based on pilot study results, their use/prevalence in prior literature, and their differences in design.

The gaze line cue consists of a line drawn from the user's FOV to the target location (see Fig.3.1a). This allows it to convey both direction (by following the line) and location. During both of the pilot studies the gaze line cue produced favorable results and was preferred by participants. Additionally, a prior study utilized this cue and demonstrated favorable outcomes for identifying buildings in a simulated JTAC task [111].

Several prior studies have also produced favorable results with the 2D wedge design [72, 112]. This cue overlays a triangle on a 2D UI plane. The tip of the triangle is situated over the target location and the triangle scales as the center of the user's FOV approaches the target location (see Fig.3.1b). This allows the 2D wedge to convey both direction and location, similar to the gaze line. Additionally, the 2D wedge produced strong results in both of the pilot studies, however it was not as preferred as the gaze line.

Unlike the gaze line and 2D wedge, the 3D arrow does not provide location information, only direction. The 3D arrow cue consists of an arrow model that rotates such that the tip of the arrow points towards the target (see Fig.3.1c).This in turn, requires that the user extrapolates the target location based on the direction indicated by the arrow. Several variations of the 3D arrow cue exist, but all typically include a 3D model that points towards a target [29, 38, 91, 126]. The inclusion of the 3D arrow in this study is due to it's prevalence in prior literature, as well as the lack of location information when compared to the gaze line and 2D wedge cues.

3.1.2 Environment, Target, and Distractors

The search environment was rendered over an ArCuo marker, and was presented as a 8 by 4 grid of windows with a black background rendered behind (see Fig. 3.2). During a search 20 of these windows would be filled (out of 32). One of the windows would contain the target. The other 19 filled locations would contain the 19 different distractors. The target was the figure wearing a military helmet and vest (see Fig. 3.3). Each distractor was created to have varying degrees of similarity to the target stimuli.

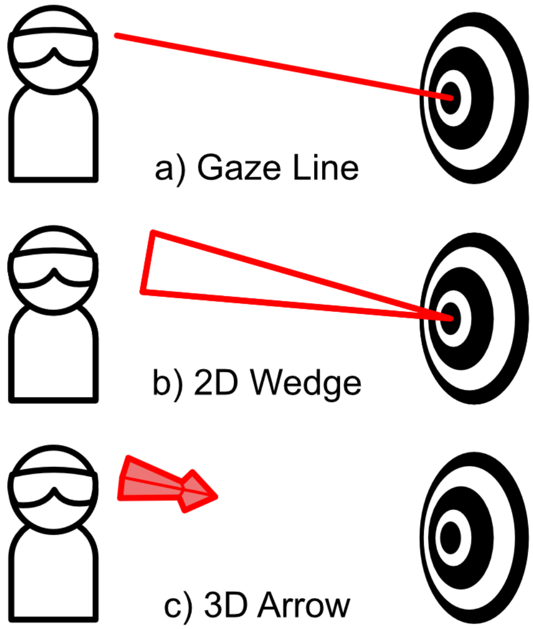

Figure 3.1: The different cue designs used for this experiment. The gaze line (a) consists of a single line drawn from the user's FOV to the target location in 3D space. The 2D wedge (b) uses a scaling triangle one of the vertices centered over a 2D translation of the target location from 3D space. The 3D arrow (c) consists of a 3D arrow model fixed to the user's FOV that rotates to always point toward the target location in 3D space.

Figure 3.2: The search field used for this experiment. This image would be overlaid onto an ArCuo marker.

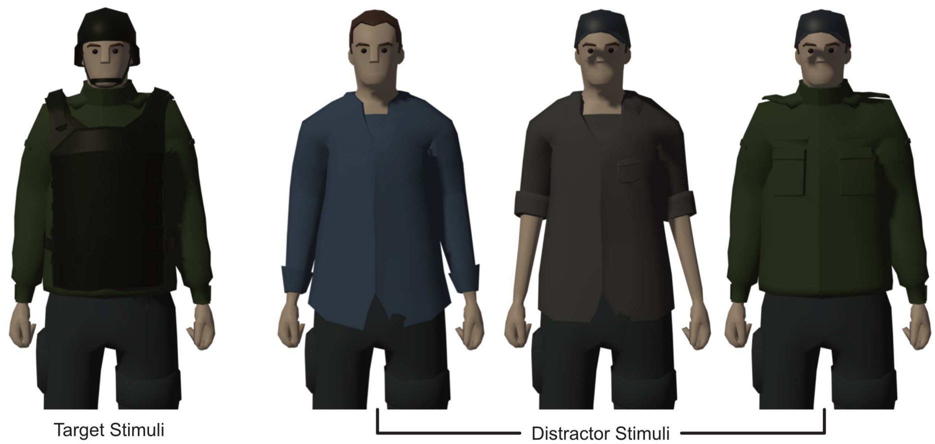

Figure 3.3: The target stimuli (left) and a selection of distractor stimuli (right 3) with varying similarity to the target.

3.2 Experiment Design

This study utilized a within subjects design with 4 different conditions (no cue, gaze line, 2D wedge, and 3D arrow). The order that a participant was exposed to these cues was counter balanced using Latin Squares. A total of 64 different search configurations were pre-loaded into the application using a configuration text file; this allowed the target stimuli to appear in every window twice. This file contained each configuration on a separate line. Windows were encoded with a number to denote their location. Each line contained numbers to denote the window location of the target and the 19 distractors. The order of the 64 configurations was randomly shuffled for each participant and between each cue condition using a random without replacement algorithm.

3.2.1 Variables

During the experiment time, accuracy, and head rotation metrics were automatically collected during each search. Time was collected by adding the time between each frame, during each search, until the participant made a selection and automatically moved on to the next search configuration. Accuracy was indicated by a 0 or 1 depending on if the target was correctly selected or not (1 indicated a correct selection). Head rotation added the total change in degrees between each frame, until the participant made a selection, and automatically moved on to the next search configuration.

In addition to time, accuracy, and head rotation, subjective measures were also collected. The overall reported workload was collected using the NASA TLX shortened questionnaire [130, 131]. Additionally, each participant was asked to rank each of the cue conditions from most to least preferred.

Table 3.1: Dependant variables and the measurement method for each.

Variables and Measures

Variable	Measure
Time	Automatically measured. Additive times since current search configuration began.
Accuracy	Automatically measured. Coded as 0 if wrong window was selected. Coded as 1 if correct window was selected.
Head Rotation	Automatically measured. Additive rotation since current search configuration began.
NASA TLX	Collected after AR portion of study using online questionnaire.
Cue Ranking	Collected after AR portion of study using online questionnaire.

3.3 Procedure

After consenting to participant, participants completed a demographics survey. This survey also included questions on prior experience with AR and VR HMDs. Once complete participants moved on to the AR portion of the study. They stood 2.5 meters away from a blank wall containing only an ArCuo marker (for registration and spawning of the virtual search environment, distractors, and target). Once the HMD identified the marker, participants completed an orientation in the HMD to become familiar with the search task, the target stimuli, and the controls (in the form of a virtual slide show). They were allowed to engage with the orientation material for as long as they wanted.

Once they felt ready they began the AR search portion of the experiment. Each participant used all four cue conditions (no cue, gaze line, 2D wedge, 3D arrow). The order of these cues was counter balanced via Latin Squares to minimize any order or learning effects. Each participant completed 64 searches with each of the cue conditions. When using a cue (anything other than the no cue condition), the cue always cued the target location. The searches were controlled by a configuration file, and the order of the searches was shuffled using a random without replacement algorithm between each participant and between each cue condition. This produced a total of 256 searches per participant (64 search configurations x 4 cue conditions). Between each condition the participants were granted a 20 second break.

After completing all of the searches, participants completed a post experiment survey consisting of the NASA TLX questionnaire and a cue ranking question. Additional subjective feedback was solicited regarding their experience. They were then thanked for their participation. Each participant was compensated with either a $25 Amazon E-Gift card or extra credit in an applicable course.

3.4 Participants

A total of 14 participants were recruited for this experiment. Of theses 8 identified as male, 5 as female, and 1 as non binary, with an average age of 22.79. Six participants reported prior use of AR, however most of this use was reported as demonstrations, part of a tour, or as part of another AR related study. Six participants also reported prior use of VR. On average participants used a computer for 45.23 hours each week (both work and pleasure).

Chapter 4

Results and Analysis

4.1 Results

Data was collected to examine the response time in seconds, the response accuracy, the participant's head rotation, subjective demand (via the NASA TLX), and cue preference. Results heavily favor the gaze line cue, although any cue improved upon the base line no cue condition. Data was checked for normal distribution using QQ-plots.

4.1.1 Time

A one way ANOVA was conducted, and indicated a significant result for search time ($F(3, 42) = 56.65, p < 0.001$). Post hoc analysis with Bonferroni correction indicated significance between each of the condition pairs for search time (see Figure: 4.1). Effect size analysis using Cohen's D revealed a large effect size for all of the condition pairs (see Table 4.1).

The fastest search time was produced by the gaze line cue (1.35 seconds). This was nearly 5 seconds less than the no cue condition (6.16 seconds). The 2D wedge cue produced the second fastest response time at 1.88 seconds. While the 3D arrow was still faster than the no cue condition by 3.29 seconds it produced the slowest response time of the cued conditions (2.87 seconds).

Table 4.1: Search time effect size (Cohen's D) for each condition pair.

Search Time Effect Size

Condition Pair	Cohen's D
No Cue X 3D Arrow	1.61
No Cue X Gaze Line	2.55
No Cue X 2D Wedge	2.17
3D Arrow X Gaze Line	1.8
3D Arrow X 2D Wedge	0.97
Gaze Line x 2D Wedge	0.81

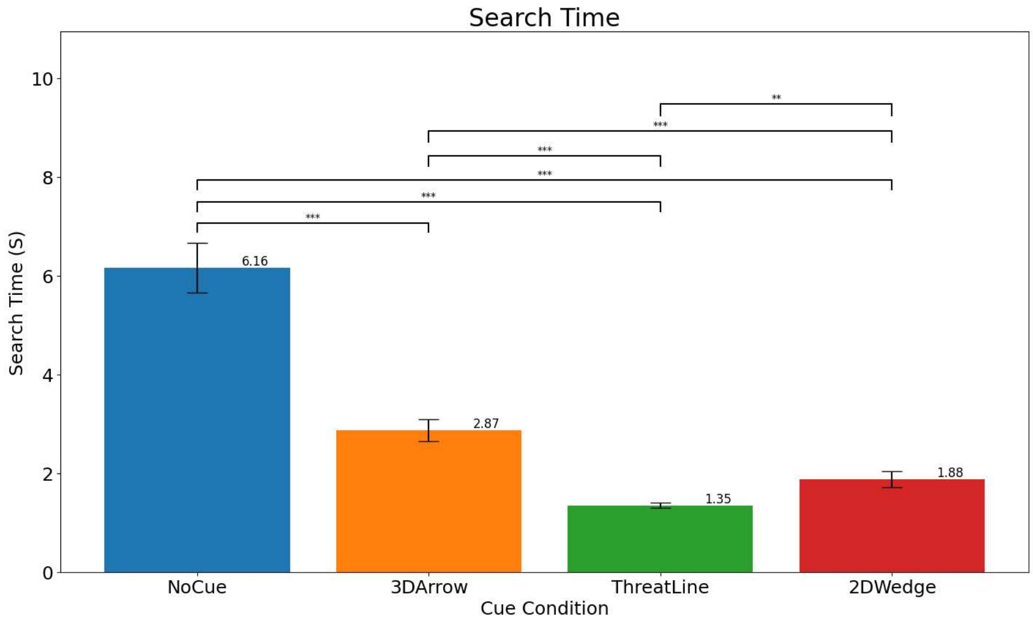

Figure 4.1: Average search time for each cue condition.

4.1.2 Accuracy

For accuracy, one participant's dataset was identified as an outlier and excluded. Anova results for the remaining data points revealed a significant effect for accuracy ($F(3, 39) = 4.89, p < 0.01$). Post hoc analysis revealed a significant effect between the no cue and gaze line condition ($p < 0.05$), for the no cue and 2D wedge condition ($p < 0.01$), for the 3D arrow and gaze line condition ($p < 0.05$), and for the 3D arrow and 2D wedge condition ($p < 0.05$). For each condition pair, the observed effect size was small (d between 0.2 and 0.49: see Table: 4.2).

As with search time the strongest accuracy results were produced by the gaze line condition with a mean accuracy of 99.11%. This was followed by the 2D wedge condition (97.43%) and then the 3D arrow condition (94.08%). Each of the cued conditions improved on the no cue condition (87.17%), with the gaze line providing an 11.94% increase in accuracy.

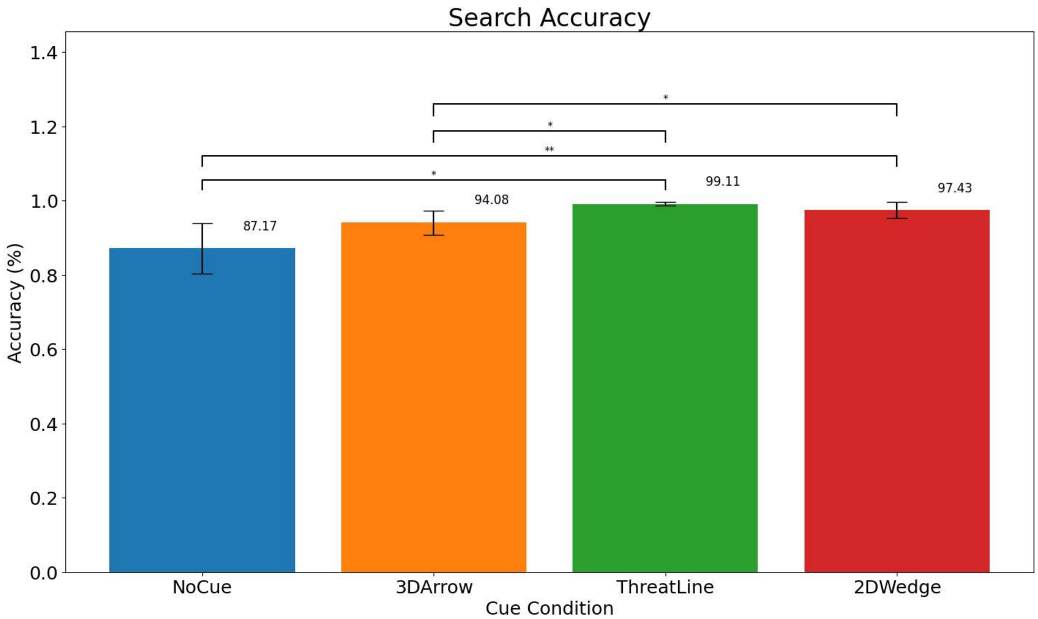

Figure 4.2: Average accuracy with each cue condition.

Table 4.2: Accuracy effect size (Cohen's D) for each condition pair.

Accuracy Effect Size

Condition Pair	Cohen's D
No Cue X 3D Arrow	0.24
No Cue X Gaze Line	0.47
No Cue X 2D Wedge	0.38
3D Arrow X Gaze Line	0.41
3D Arrow X 2D Wedge	0.23
Gaze Line x 2D Wedge	0.21

4.1.3 Head Rotation

Strong significance also emerged for the head rotation exhibited by participants though their searches ($F(3, 42) = 58.28, p < 0.001$). Post hoc analysis revealed significant effects for all condition pairs except for gaze line and 2D wedge. The interaction between the no cue condition and all cue conditions (3D arrow, gaze line, and 2D wedge) all exhibited very strong significance ($p < 0.001$). The interaction between the 3D arrow and gaze line demonstrated strong significance

($p < 0.01$), and the interaction between the 3D arrow and 2D wedge also indicated significance ($p < 0.05$).

The average head rotation per search with the no cue condition was 2696.4 degrees. Each of the cue conditions reduced this to less than 1250 degrees with the gaze line exhibiting the least amount of movement (727.36 degrees). This was followed by the 2D wedge with 868.93 degrees of rotation and then the 3D arrow with 1115.01 degrees of rotation.

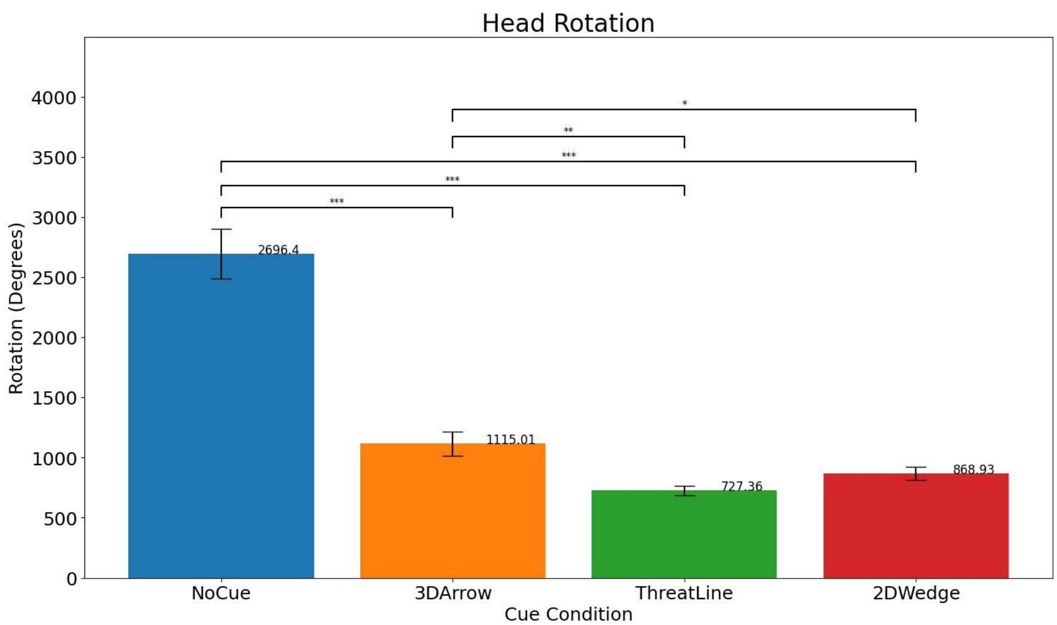

Figure 4.3: Average head rotation in degrees for each cue condition.

While each cue improved over the baseline no cue condition, the gaze line condition produced the most favorable results (see Table 4.3). The gaze line cue produced a 4.81 second faster response time with an accuracy increase of 11.94% over the baseline no cue condition. Head movement with the gaze line was also reduced by 1969.4 degrees over the no cue condition.

Table 4.3: Search time, accuracy, and head rotation means across all 4 cue conditions. The gaze line cue demonstrated the lowest search time, highest accuracy, and least amount of head rotation.

Search Time, Accuracy, and Head Rotation

Cue Cond.	**Search Time**	**Accuracy**	**Rotation**
No Cue	6.16s	87.17%	2696.4 deg.
3D Arrow	2.87s	94.08%	1115.01 deg.
Gaze Line	**1.35s**	**99.11%**	**727.36 deg.**
2D Wedge	1.88s	97.43%	868.93 deg.

4.1.4 NASA TLX and Cue Rankings

When asked to rank cue preferences from 1 (highest) to 4 (least) the gaze line cue received the most favorable rankings. The gaze line cue had 10 participants rank it as their preferred cue. The second most preferred was tied between the 3D arrow and the 2D wedge with each having 5 votes for rank 2. The 2D wedge received the most votes for rank 3. The no cue condition was the lowest ranked with 10 participants ranking the condition last.

A favorability score was also calculated using these rankings (see Table 4.4). For each rank 1 rating the cue was granted 4 points, for each rank 2 rating the cue was granted 3 points, for each rank 3 rating the cue was granted 2 points, and for each rank 4 rating the cue was granted 1 point. The gaze line had the highest score with 45 points followed by the 3D arrow with 32 points, then the 2D wedge with 28 points, and finally the no cue condition with 15 points.

Table 4.4: Number of times each cue appeared at different ranking (1 is best 4 is worst) with associated favorability scores.

Cue Ranking

Rank	No Cue	Gaze Line	3D Arrow	2D Wedge
1	0	**10**	2	0
2	1	1	**5**	**5**
3	1	1	4	**6**
4	**10**	0	1	1
Score	15	**45**	32	28

NASA TLX scores also reveal a preference for the gaze line cue (see Table 4.5). For mental demand the gaze line was rated as 2 (on a scale of 1-7). Physical demand was also rated low for the gaze line cue (1.57 out of 7). Participants also reported the highest feeling of success with the gaze line cue (6.79 out of 7). The total amount of work required to complete the task was also lowest with the gaze line cue (1.86 out of 7) as was the level of reported annoyance (1.43 out of 7). However, the condition with the most favorable pacing result was the 3D arrow cue (3.07 out of 7) with the gaze line and no cue condition both having the least favorable scores at 3.36 (out of 7).

Table 4.5: Average NASA TLX score for each category and cue.

NASA TLX Scores

TLX Measure	No Cue	Gaze Line	3D Arrow	2D Wedge
Mental Demand	5.5	**2**	4.14	3.57
Physical Demand	3.86	**1.57**	3	2.42
Pacing	3.36	3.36	**3.07**	3.14
Success	4.21	**6.79**	5.71	5.57
Total Work	5.5	**1.86**	3.64	3.14
Annoyance	4	**1.43**	2.93	3

4.1.5 Hypothesis Testing

Seven hypothesis were formulated by reviewing prior literature:

- H1: Any cue condition will improve upon the baseline condition.
- H2: Cues that take up more of the user's FOV will be less effective.
- H3: The more salient the cue design the more effective it will be.
- H4: Cues with more embedded information will produce faster search times.
- H5: Cues with more embedded information will produce higher accuracy rates.
- H6: Cues with more embedded information will reduce user movement.

- H7: Cues with more embedded information will reduce task demand.

The significant decrease in search time, increase in accuracy, and reduction in head movement using any of the three cue conditions over the baseline no cue condition supports H1. Each cue produced a more favorable result over the base line condition with the exception of the 3D arrow accuracy.

There is partial support for H2, as the 2D wedge did consume more screen space than the gaze line and the gaze line produced stronger results, however, the 3D arrow was not as invasive as the 2D wedge, but performed worse. This may be due to the embedded position information that was not present in the 3D arrow, but was present in the 2D wedge design.

Both the 2D wedge and the gaze line had a brighter contrast and more dynamic presentation than the 3D arrow. As such they would be classified as more salient than the 3D arrow. Due to this the cue performance results support H3 as the 3D arrow would be considered the least salient and produced the worst search time, accuracy, and head rotation results of the three cue designs.

The 3D arrow did not have location information embedded into the design. Instead, participants had to extrapolate position based on the direction the 3D arrow was indicating. As such the gaze line and 2D wedge contained more embedded information (direction and position). Due to this H4 through H6 are all supported as the gaze line and 2D wedge produced the faster search times (H4), highest accuracy (H5), and the least amount of head rotation (H6). The differences between the gaze line and wedge may be related to the mental effort and intuitiveness of the design, resulting in the gaze line's better performance.

The last hypothesis (H7) is partially supported by the NASA TLX results. The gaze line and 2D wedge (which have embedded location information not present in the 3D arrow design) produced the lowest mental demand, physical demand, and total work load scores. However, other scores such as the perceived pacing and the reported success did not follow the trend favoring the gaze line and the 2D wedge.

Based on these results H1 (any cue condition will improve upon the baseline condition) is supported. H2 (cues that take up more of the user's FOV will be less effective) was partially supported, but H3 (the more salient the cue design the more effective it will be, as long as the cue is completely accurate) was supported. H4 (cues with more embedded information will produce faster search times), H5 (cues with more embedded information will produce higher accuracy rates) and H6 (cues with more embedded information will reduce user movement) are all supported, but H7 (cues with more embedded information will reduce task demand) is only partially supported.

4.2 Discussion

While any cue increased search efficacy over the baseline no cue condition, the gaze line condition provided for the strongest results. Using the gaze line condition decreased search time by nearly 5 seconds and increased accuracy by almost 12%. Head rotation was also reduced by almost 2000 degrees or nearly 5.5 full 360 degree rotations. This is in line with other studies such as Mifsud et al. [111], that also found the gaze line cue to be particularly effective.

The 2D wedge provided for the second strongest results with a reduction of 4.28 seconds for search time and an increase of about 10% in accuracy. As with the gaze line cue the 2D wedge also shows both the direction to and location of the target [72, 112]. However, the reduced performance from the 2D wedge in contrast to the gaze line may be due to a cluttering effect [70, 71] as the 2D wedge scales to be much larger, and thus occupies more screen realestate, than the gaze line. This effect may also be due to an increase in cognitive resources required to understand the 2D wedge design as the wedge may be more conceptually new than the gaze line. Lines and line motifs are often used for guidance in daily life, whether it be the center lines of a road to help guide a car forward, or a laser pointer to help guide attention on a slide show. As such less cognitive resources may be required to process and understand the information embedded in the 2D wedge design.

Despite the prevalence of the 3D arrow cue in literature [29, 38, 91, 126], the 3D arrow performance provided for the smallest benefit over the baseline no cue condition. This is most likely due to its lack of embedded position information. Unlike both the gaze line and 2D wedge, the 3D arrow only communicates direction information to the user. As such, extrapolation of the position based on direction is required. This would not only increase cognitive load, but would also

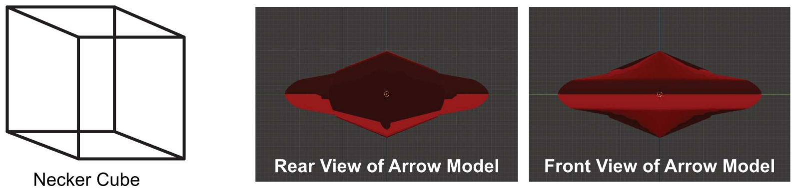

Figure 4.4: Left: the Necker Cube [7, 8]. Middle: rear view of the arrow model. Right: front view of the arrow model. The two views of the arrow model have similar shapes resulting in some level of ambiguity.

increase processing time (thus increasing search time), and may be more error prone due to ambiguity issues. The problem of 3D ambiguity was originally discussed in conjunction with the *Necker Cube* [7, 8]. First published in 1832 the Necker Cube is a wireframe 3D cube that, through ambiguity of the viewing angle, causes the perceived front face to switch leading to multiple possible perceptions of the same drawing (see Fig.4.4). This effect may be in play with the 3D arrow cue, especially when viewed either head on or directly behind as it can be difficult to discern it's orientation; thus increasing the challenge of extrapolating position from direction.

The benefits of the gaze line cue were supported further by both the NASA TLX scores and the cue ranking results. With a favorability score of 45 the gaze line was ranked as the preferred cue with 10 out of 12 participants ranking it a their favorite. This was followed by the 3D arrow, then by the 2D wedge (see Table 4.4), despite the 2D wedge producing more favorable results. This may be further evidence of the conceptual complexity of the 2D wedge over the other cues, as well as evidence for the benefit of embedding both position and direction information into cue designs. Despite the lack of preference for the 2D wedge over the 3D arrow it still provided for better search time and accuracy results than the 3D arrow.

NASA TLX results further solidify the gaze line cue as the strongest condition in this study. The gaze line cue greatly reduced the reported mental demand, physical demand, and total perceived workload of the task. This supports the cue as a strong choice for reducing task workload and provides explanation for the improved search results with the gaze line. The higher mental demand for the 3D arrow continues to indicated that the potential ambiguity (*a la* Necker Cube [7, 8]) and lack of position information did increase the mental effort required to use the 3D arrow over the other cue conditions. The one score where the 3D arrow was more favorable than the other conditions was in reported pacing. This is a measure of how rushed the participant felt with higher scores being unfavorable. All of the conditions were within a half point of each other, however the 3D arrow was reported as the least rushed. Conversely, the gaze line was tied with the no cue condition as the most rushed. This may indicated that the more readily available information is the more enticed participants feel to act upon the information. This in turn may lead to automation bias issues when errors are introduced into a system [76–79]. Participants also reported feeling the most successful with the gaze line cue with an average score nearing perfect. However, this may also indicate the potential for automation bias to be encountered when errors are introduced, as the participants may be over confident with the system recommendations.

Based on these results any cue would seem to be beneficial for visual search tasks, however not all cues are equal. The gaze line is a particularly strong choice for visual search tasks, however this may only be true in cases such as this study, where the cue is always correct. There is also the potential danger of reducing the mental resource demand too much, which may lead to performance issues by falling out of the "Goldilocks zone" [20, 39]. As such cues should be vetted and tested to determine benefit for any given task.

4.2.1 Limitations and Future Directions

Results from this study point to potential issues with automation bias when errors are included, however this study did not introduce any errors into the work. When using cues the cue always pointed to the correct location, however this would not be the case with a live cueing system driven by computer vision or machine learning (ML) techniques [75]. Due to the nature of ML detection errors would inevitably arise. These errors, in combination with machine bias issues [76–79], can lead the best cues in a perfect condition to suffer greatly and become worse options for cueing [25, 74, 75].

This study was also limited to a forward facing search field. Due to this, the ambiguity effects observed by the 3D arrow may have been minimized in this study. Furthermore, the other cueing conditions did not have to account for targets being in front of or behind the user which may lead to different results, as other unconsidered issues with ambiguity or direction indication may exist with the gaze line or 2D wedge condition.

A lab setting was used for this study which is not representative of real world AR use. Aspects such as occlusion [91, 108] and lighting [28, 95] may cause issues with optical see through rendering used in HMDs such as the Magic Leap 2. Environmental clutter may also be a factor that limits the usefulness of the cue design presented [115]. For instance, a room with many brightly colored red toys may overcome the red gaze line rendering limiting its functionality.

Only a single cue was present at any given time, however there are often situations where an individual must pay attention to multiple targets simultaneously. Increasing the amount of cues present may have adverse affects due to clutter [70, 71]. As such, determining the maximum number of cues that can be simultaneously utilized is necessary before including more than one cue in a task. This number may also differ from cue to cue due to the differences in mental effort required to process and attend to those cues. For instance, three gaze line cues may be far easier to attend to than three 2D wedge cues.

Each of these suggests future research topics for exploring aspects of visual cueing and may lead to new cue designs and innovations.

Chapter 5

Conclusion

With the increasing attention for using AR for visual cueing it is important that cue designs are evaluated and compared. This study presents a within-subjects, four factored experimental design, where participants use four cue conditions (no cue, gaze line, 2D wedge, and 3D arrow) to search for target stimuli in a forward facing AR environment. Results suggest that any cue intervention is beneficial over the base line no cue condition, however the gaze line cue provided for the fastest search time, the highest search accuracy, the least amount of head movement, reduced mental and physical demand, and was preferred by participants. Additionally, results support embedding both the direction to, and position of the target stimuli in cue designs, as this may reduce ambiguity and help to improve results. However, over complicated designs may require more mental effort to process. Furthermore, there is a potential danger with too salient designs (i.e. lots of motion, high contrast, large size, etc.), as the reduction in cognitive resources required to complete the task may in turn lead to a reduction of performance through automation bias and falling out of the "Goldilocks zone."

Bibliography

[1] Walter Schneider and Richard M. Shiffrin. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84(1):1–66, January 1977.

[2] Anne M. Treisman and Garry Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12(1):97–136, January 1980.

[3] R.C. Atkinson and R.M. Shiffrin. Human Memory: A Proposed System and its Control Processes. In *Psychology of Learning and Motivation*, volume 2, pages 89–195. Elsevier, 1968.

[4] G. J. Hitch and A. D. Baddeley. Verbal Reasoning and Working Memory. *Quarterly Journal of Experimental Psychology*, 28(4):603–621, November 1976.

[5] Paul Milgram and Fumio Kishino. A TAXONOMY OF MIXED REALITY VISUAL DISPLAYS. *IEICE Transactions on Information Systems*, E77-D(12):15, 1994.

[6] Nicholas Katzakis, Jonathan Tong, Oscar Ariza, Lihan Chen, Gudrun Klinker, Brigitte Röder, and Frank Steinicke. Stylo and handifact: modulating haptic perception through visualizations for posture training in augmented reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, pages 58–67, Brighton United Kingdom, October 2017. ACM.

[7] Jürgen Kornmeier and Michael Bach. The Necker cube—an ambiguous figure disambiguated in early visual processing. *Vision Research*, 45(8):955–960, April 2005.

[8] Mareike Wilson, Lukas Hecker, Ellen Joos, Ad Aertsen, Ludger Tebartz Van Elst, and Jürgen Kornmeier. Spontaneous Necker-cube reversals may not be that spontaneous. *Frontiers in Human Neuroscience*, 17:1179081, May 2023.

[9] Brendan Kelley, Anil Ufuk Batmaz, Michael Humphrey, Cyane Tornatzky, Rosa Martey, and Francisco Ortega. Emergent Individual Factors for AR Education and Training. In George Bebis, Golnaz Ghiasi, Yi Fang, Andrei Sharf, Yue Dong, Chris Weaver, Zhicheng Leo, Joseph J. LaViola, and Luv Kohli, editors, *Advances in Visual Computing*, volume 14362, pages 27–38. Springer Nature Switzerland, Cham, 2023. Series Title: Lecture Notes in Computer Science.

[10] Clemens Hoffmann, Sebastian Büttner, Michael Prilla, and Kai Wundram. Impact of augmented reality guidance for car repairs on novice users of AR: a field experiment on familiar and unfamiliar tasks. In *Proceedings of the Conference on Mensch und Computer*, pages 279–289, Magdeburg Germany, September 2020. ACM.

[11] Radha Kumaran, You-Jin Kim, Anne E Milner, Tom Bullock, Barry Giesbrecht, and Tobias 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*, pages 1–17, Hamburg Germany, April 2023. ACM.

[12] Nicola Binetti, Luyan Wu, Shiping Chen, Ernst Kruijff, Simon Julier, and Duncan P. Brumby. Using visual and auditory cues to locate out-of-view objects in head-mounted augmented reality. *Displays*, 69:102032, September 2021.

[13] Yuhang Zhao, Elizabeth Kupferstein, Brenda Veronica Castro, Steven Feiner, and Shiri 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*, pages 387–402, New Orleans LA USA, October 2019. ACM.

[14] Uwe Gruenefeld, Dag Ennenga, Abdallah El Ali, Wilko Heuten, and Susanne 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*, pages 109–118, Brighton United Kingdom, October 2017. ACM.

[15] Jen-Shuo Liu, Portia Wang, Barbara Tversky, and Steven Feiner. Adaptive Visual Cues for Guiding a Bimanual Unordered Task in Virtual Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 431–440, Singapore, Singapore, October 2022. IEEE.

[16] Kody R. Dillman. *A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality*. Dissertation, University of Calgary, Calgary, Canada, September 2018.

[17] Xinyu Hu, Alec G. Moore, James Coleman Eubanks, Afham Ahmed Aiyaz, and Ryan 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)*, pages 63–69, Atlanta, GA, USA, March 2020. IEEE.

[18] Xinyu Hu, Alec Moore, James Coleman Eubanks, Afham Aiyaz, and Ryan P. McMahan. Evaluating Interaction Cue Purpose and Timing for Learning and Retaining Virtual Reality Training. In *Symposium on Spatial User Interaction*, pages 1–9, Virtual Event Canada, October 2020. ACM.

[19] Dieter Schmalstieg and Tobias Höllerer. *Augmented reality: principles and practice*. Addison-Wesley usability and HCI series Always Learning. Addison-Wesley, Boston Columbus Indianapolis, 2016.

[20] Christopher D. Wickens, Robert S. Gutzwiller, and Jason S. McCarley. *Applied attention theory*. CRC Press, Taylor & Francis Group, Boca Raton, second edition edition, 2023.

[21] Christopher D. Wickens, William S. Helton, Justin G. Hollands, and Simon Banbury. *Engineering psychology and human performance*. Routledge, New York (N.Y.), fifth edition edition, 2022.

[22] Unity Technologies. Unity, 2022.

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

[24] M. P. Eckstein. Visual search: A retrospective. *Journal of Vision*, 11(5):14–14, December 2011.

[25] Amelia C. Warden, Christopher D. Wickens, Domenick Mifsud, Shannon Ourada, Benjamin A. Clegg, and Francisco R. Ortega. Visual Search in Augmented Reality: Effect of Target Cue Type and Location. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1):373–377, September 2022.

[26] Uwe Gruenefeld, Lars Prädel, and Wilko Heuten. Locating nearby physical objects in augmented reality. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, pages 1–10, Pisa Italy, November 2019. ACM.

[27] Alexander Marquardt, Christina Trepkowski, Tom David Eibich, Jens Maiero, Ernst Kruijff, and Johannes Schoning. Comparing Non-Visual and Visual Guidance Methods for Narrow Field of View Augmented Reality Displays. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3389–3401, December 2020.

[28] You-Jin Kim, Radha Kumaran, Ehsan Sayyad, Anne Milner, Tom Bullock, Barry Giesbrecht, and Tobias Hollerer. Investigating Search Among Physical and Virtual Objects Under Different Lighting Conditions. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3788–3798, November 2022.

[29] Uwe Gruenefeld, Daniel Lange, Lasse Hammer, Susanne Boll, and Wilko Heuten. FlyingARrow: Pointing Towards Out-of-View Objects on Augmented Reality Devices. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays*, pages 1–6, Munich Germany, June 2018. ACM.

[30] Uwe Gruenefeld, Ilja Koethe, Daniel Lange, Sebastian Weiss, and Wilko Heuten. Comparing Techniques for Visualizing Moving Out-of-View Objects in Head-mounted Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 742–746, Osaka, Japan, March 2019. IEEE.

[31] Joel D Miller, Dr Martine Godfroy-Cooper, and Zoltan P Szoboszlay. Augmented-Reality Multimodal Cueing for Obstacle Awareness: Towards a New Topology for Threat-Level Presentation. page 32, 2019.

[32] Ane San Martín and Johan Kildal. Audio-visual AR to Improve Awareness of Hazard Zones Around Robots. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, Glasgow Scotland Uk, May 2019. ACM.

[33] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. The Effects of Sharing Awareness Cues in Collaborative Mixed Reality. *Frontiers in Robotics and AI*, 6:5, February 2019.

[34] Seungwon Kim, Gun Lee, Weidong Huang, Hayun Kim, Woontack Woo, and Mark Billinghurst. Evaluating the Combination of Visual Communication Cues for HMD-based Mixed Reality Remote Collaboration. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, Glasgow Scotland Uk, May 2019. ACM.

[35] Patrick Perea, Denis Morand, and Laurence Nigay. Spotlight on Off-Screen Points of Interest in Handheld Augmented Reality: Halo-based techniques. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pages 43–54, Daejeon Republic of Korea, November 2019. ACM.

[36] Patrick Perea, Denis Morand, and Laurence Nigay. Halo3D: a technique for visualizing off-screen points of interest in mobile augmented reality. In *Proceedings of the 29th Conference on l'Interaction Homme-Machine*, pages 43–51, Poitiers France, August 2017. ACM.

[37] Andrew Yoshimura, Adil Khokhar, and Christoph W. Borst. Eye-gaze-triggered Visual Cues to Restore Attention in Educational VR. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1255–1256, Osaka, Japan, March 2019. IEEE.

[38] Nilotpal Biswas, Arpit Singh, and Samit Bhattacharya. AroundArrow: Off-Screen POI visualization for handheld Augmented Reality in vertically dense regions. In *2022 IEEE In-*

ternational Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pages 570–575, Singapore, Singapore, October 2022. IEEE.

[39] Christopher Wickens. Attention: Theory, Principles, Models and Applications. *International Journal of Human–Computer Interaction*, 37(5):403–417, March 2021.

[40] Jeremy M. Wolfe and Todd S. Horowitz. Five factors that guide attention in visual search. *Nature Human Behaviour*, 1(3):0058, March 2017.

[41] Jeremy M. Wolfe. Visual search. *Current Biology*, 20(8):R346–R349, April 2010.

[42] Wieske Van Zoest and Mieke Donk. Bottom-up and Top-down Control in Visual Search. *Perception*, 33(8):927–937, August 2004.

[43] Jan Theeuwes. Top–down and bottom–up control of visual selection. *Acta Psychologica*, 135(2):77–99, October 2010.

[44] Brian A. Anderson, Patryk A. Laurent, and Steven Yantis. Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, 108(25):10367–10371, June 2011.

[45] Melissa Le-Hoa Võ and Jeremy M. Wolfe. The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, 1339(1):72–81, March 2015.

[46] Mark B. Neider and Gregory J. Zelinsky. Scene context guides eye movements during visual search. *Vision Research*, 46(5):614–621, March 2006.

[47] Monica S. Castelhano and John M. Henderson. Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 33(4):753–763, 2007.

[48] Jeremy M. Wolfe. Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28(4):1060–1092, August 2021.

[49] Sungwon Chung and Johnny V. Sparks. Motivated Processing of Peripheral Advertising Information in Video Games. *Communication Research*, 43(4):518–541, June 2016.

[50] Jacob T Fisher and René Weber. The Limited Capacity Model of Motivated Mediated Message Processing (LC4MP). page 11, 2020.

[51] Wolfgang Schnotz and Christian Kürschner. A Reconsideration of Cognitive Load Theory. *Educational Psychology Review*, 19(4):469–508, October 2007.

[52] John Sweller and Paul Chandler. Evidence for Cognitive Load Theory. *Cognition and Instruction*, 8(4):351–362, December 1991.

[53] Alan D. Baddeley, Michael W. Eysenck, and Michael C. Anderson. *Memory*. A Psychology Press book. Routledge, Taylor & Francis Group, London New York, third edition edition, 2020.

[54] Alan Baddeley. Working Memory. *Science*, 255(5044):556–559, January 1992.

[55] Alan Baddeley. Working Memory. In *Working Memory and Education*, pages 1–31. Elsevier, 2006.

[56] Wen Jia Chai, Aini Ismafairus Abd Hamid, and Jafri Malin Abdullah. Working Memory From the Psychological and Neurosciences Perspectives: A Review. *Frontiers in Psychology*, 9:401, March 2018.

[57] John F. Kihlstrom and Lillian Park. Cognitive Psychology: Overview. In *Reference Module in Neuroscience and Biobehavioral Psychology*, page B9780128093245217021. Elsevier, 2018.

[58] Steven J. Luck and Michelle A. Ford. On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences*, 95(3):825–830, February 1998.

[59] Nelson Cowan. Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, 26(2):197–223, June 2014.

[60] Paul M. Bays and Masud Husain. Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science*, 321(5890):851–854, August 2008.

[61] Wei Ji Ma, Masud Husain, and Paul M Bays. Changing concepts of working memory. *Nature Neuroscience*, 17(3):347–356, March 2014.

[62] Annie Lang. The Limited Capacity Model of Mediated Message Processing. *Journal of Communication*, 50(1):46–70, March 2000.

[63] Jeff Pankin. Schema Theory, 2013.

[64] Kalianne L. Neumann and Theodore J. Kopcha. The Use of Schema Theory in Learning, Design, and Technology. *TechTrends*, 62(5):429–431, September 2018.

[65] Mary B. McVee, Kailonnie Dunsmore, and James R. Gavelek. Schema Theory Revisited. *Review of Educational Research*, 75(4):531–566, December 2005.

[66] Fausto Giunchiglia, Pavel Shvaiko, and Mikalai Yatskevich. Semantic Schema Matching. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Robert Meersman, and Zahir Tari, editors, *On the Move to Meaningful Internet Systems 2005: CoopIS, DOA, and ODBASE*, volume 3760, pages 347–365. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005. Series Title: Lecture Notes in Computer Science.

[67] Diane Lynn Damos, editor. *Multiple-task performance*. CRC Press, Boca Raton, first edition edition, 2020. OCLC: 1178935872.

[68] Donald L. Parks and George P. Boucek. Workload Prediction, Diagnosis, and Continuing Challenges. In Grant R. McMillan, David Beevis, Eduardo Salas, Michael H. Strub, Robert Sutton, and Leo Van Breda, editors, *Applications of Human Performance Models to System Design*, pages 47–63. Springer US, Boston, MA, 1989.

[69] Bruce K. Britton and Abraham Tesser. Effects of prior knowledge on use of cognitive capacity in three complex cognitive tasks. *Journal of Verbal Learning and Verbal Behavior*, 21(4):421–436, August 1982.

[70] Nadine Moacdieh and Nadine 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, February 2015.

[71] Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. Measuring visual clutter. *Journal of Vision*, 7(2):17, August 2007.

[72] Sean Gustafson, Patrick Baudisch, Carl Gutwin, and Pourang Irani. Wedge: clutter-free visualization of off-screen locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 787–796, Florence Italy, April 2008. ACM.

[73] Christina Trepkowski, Alexander Marquardt, Tom David Eibich, Yusuke Shikanai, Jens Maiero, Kiyoshi Kiyokawa, Ernst Kruijff, Johannes Schoning, and Peter Konig. Multisensory Proximity and Transition Cues for Improving Target Awareness in Narrow Field of View Augmented Reality Displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(2):1342–1362, February 2022.

[74] Amelia C. Warden, Christopher D. Wickens, Daniel Rehberg, Benjamin A. Clegg, and Francisco R. Ortega. Information Access Effort: Are Head Movements "Cheap" or Even "Free"? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1):2203–2207, September 2022.

[75] Aditya Raikwar, Domenick Mifsud, Christopher D. Wickens, Anil Ufuk Batmaz, Amelia C. Warden, Brendan Kelley, Benjamin A. Clegg, and Francisco 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*, pages 1–9, 2024.

[76] Linda J. Skitka, Kathleen L. Mosier, and Mark Burdick. Does automation bias decision-making? *International Journal of Human-Computer Studies*, 51(5):991–1006, November 1999.

[77] Linda J. Skitka, Kathleen L. Mosier, Mark Burdick, and Bonnie Rosenblatt. Automation Bias and Errors: Are Crews Better Than Individuals? *The International Journal of Aviation Psychology*, 10(1):85–97, January 2000.

[78] Mary Cummings. Automation Bias in Intelligent Time Critical Decision Support Systems. In *AIAA 1st Intelligent Systems Technical Conference*, Chicago, Illinois, September 2004. American Institute of Aeronautics and Astronautics.

[79] Kathleen L Mosier, Linda J Skitka, Susan Heers, and Mark Burdick. Automation bias: Decision making and performance in high-tech cockpits. In *Decision Making in Aviation*, pages 271–288. Routledge, 2017.

[80] Keith E Stanovich. Why humans are cognitive misers and what it means for the great rationality debate. In *Routledge handbook of bounded rationality*, pages 196–206. Routledge, 2020.

[81] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman, and Ivan Poupyrev. *3D user interfaces: theory and practice*. Pearson always learning. Addison-Wesley, Boston, Mass. Columbus, Ohio Indianapolis New York San Francisco Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto Delhi Mexico City São Paulo Sidney Hong Kong Seoul Singapore Taipei Tokyo, second edition edition, 2017.

[82] Kody R. Dillman, Terrance Tin Hoi Mok, Anthony Tang, Lora Oehlberg, and Alex 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*, pages 1–12, Montreal QC Canada, April 2018. ACM.

[83] Xiaohe Qiu, Zhen Yang, Jinxing Yang, Qijun Wang, and Duming Wang. Impact of AR Navigation Display Methods on Wayfinding Performance and Spatial Knowledge Acquisition. *International Journal of Human–Computer Interaction*, pages 1–21, January 2023.

[84] Yuhang Zhao, Elizabeth Kupferstein, Hathaitorn Rojnirun, Leah Findlater, and Shiri Azenkot. The Effectiveness of Visual and Audio Wayfinding Guidance on Smartglasses for People with Low Vision. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, Honolulu HI USA, April 2020. ACM.

[85] Jared Van Dam, Alexander Krasne, and Joseph L. Gabbard. Drone-based Augmented Reality Platform for Bridge Inspection: Effect of AR Cue Design on Visual Search Tasks. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 201–204, Atlanta, GA, USA, March 2020. IEEE.

[86] Ronald T. Azuma. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, August 1997.

[87] Jason Jerald. *The VR book: human-centered design for virtual reality*. Number 8 in ACM books. acm, Association for Computing Machinery, New York, first edition edition, 2016.

[88] Cha Lee, Scott Bonebrake, Tobias Hollerer, and Doug A. Bowman. A replication study testing the validity of AR simulation in VR for controlled experiments. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pages 203–204, Orlando, FL, USA, October 2009. IEEE.

[89] S.R. Ellis, F. Breant, B. Manges, R. Jacoby, and B.D. Adelstein. Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, pages 138–145, Albuquerque, NM, USA, 1997. IEEE Comput. Soc. Press.

[90] Cha Lee, Steffen Gauglitz, Tobias Hollerer, and Doug A. Bowman. Examining the equivalence of simulated and real AR on a visual following and identification task. In *2012 IEEE Virtual Reality (VR)*, pages 77–78, Costa Mesa, CA, USA, March 2012. IEEE.

[91] Difeng Yu, Hai-Ning Liang, Kaixuan Fan, Heng Zhang, Charles Fleming, and Konstantinos Papangelis. 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, September 2020.

[92] Jeanine K. Stefanucci, David Brickler, Hunter C. Finney, Emi Wilson, Trafton Drew, and Sarah H. Creem-Regehr. Effects of simulated augmented reality cueing in a virtual navigation task. *Frontiers in Virtual Reality*, 3:971310, September 2022.

[93] Chieko Asakawa and Hironobu Takagi. Text Entry for People with Visual Impairments. In *Text Entry Systems*, pages 305–318. Elsevier, 2007.

[94] I. Scott MacKenzie. *Human-computer interaction: an empirical research perspective*. Morgan Kaufmann is an imprint of Elsevier, Amsterdam, 1st. edition edition, 2013.

[95] Geoffrey S. Hubona, Philip N. Wheeler, Gregory W. Shirah, and Matthew Brandt. The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer-Human Interaction*, 6(3):214–242, September 1999.

[96] Patrick Baudisch and Ruth Rosenholtz. Halo: a technique for visualizing off-screen objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 481–488, Ft. Lauderdale Florida USA, April 2003. ACM.

[97] Sebastian Günther, Julian Rasch, Dominik Schön, Florian Müller, Martin Schmitz, Jan Riemann, Andrii Matviienko, and Max Mühlhäuser. Smooth as Steel Wool: Effects of Visual Stimuli on the Haptic Perception of Roughness in Virtual Reality. In *CHI Conference on*

Human Factors in Computing Systems, pages 1–17, New Orleans LA USA, April 2022. ACM.

[98] Osama Halabi and Mohammad Saleh. Augmented reality flavor: cross-modal mapping across gustation, olfaction, and vision. *Multimedia Tools and Applications*, 80(30):36423–36441, December 2021.

[99] Sanne Boesveldt and Valentina Parma. The importance of the olfactory system in human well-being, through nutrition and social behavior. *Cell and Tissue Research*, 383(1):559–567, January 2021.

[100] Nicole Tobisková, Lennart Malmsköld, and Thomas Pederson. Multimodal Augmented Reality and Subtle Guidance for Industrial Assembly – A Survey and Ideation Method. In Jessie Y. C. Chen and Gino Fragomeni, editors, *Virtual, Augmented and Mixed Reality: Applications in Education, Aviation and Industry*, volume 13318, pages 329–349. Springer International Publishing, Cham, 2022. Series Title: Lecture Notes in Computer Science.

[101] Michele Gattullo, Alessandro Evangelista, Antonio E. Uva, Michele Fiorentino, and Joseph L. Gabbard. What, How, and Why are Visual Assets Used in Industrial Augmented Reality? A Systematic Review and Classification in Maintenance, Assembly, and Training (From 1997 to 2019). *IEEE Transactions on Visualization and Computer Graphics*, 28(2):1443–1456, February 2022.

[102] Yuki Harada and Junji Ohyama. Quantitative evaluation of visual guidance effects for 360-degree directions. *Virtual Reality*, 26(2):759–770, June 2022.

[103] Stefano Burigat, Luca Chittaro, and Silvia Gabrielli. Visualizing Locations of Off-Screen Objects on Mobile Devices: A Comparative Evaluation of Three Approaches.

[104] Enricoandrea Laviola, Michele Gattullo, Alessandro Evangelista, Michele Fiorentino, and Antonio Emmanuele Uva. In-situ or side-by-side? A user study on augmented reality maintenance instructions in blind areas. *Computers in Industry*, 144:103795, January 2023.

[105] Laura Malinin, Leah Scolere, and Brendan Kelley. A living laboratory: Exploring affordances of an academic incubator as a physical-digital metaverse. In *3rd International Conference of Possibility Studies*, volume 3, Dublin, Ireland, July 2023. Possibility Studies Network.

[106] James A Walsh, James Baumeister, and Bruce H Thomas. Spatial Augmented Reality Visibility and Line-of-Sight Cues for Building Design. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pages 1–5, Osaka Japan, December 2021. ACM.

[107] Ian P. Howard. Spatial vision within egocentric and exocentric frames of reference. *NASA. Ames Research Center, Visually Guided Control of Movement*, April 1991.

[108] Niklas Osmers and Michael Prilla. Getting out of Out of Sight: Evaluation of AR Mechanisms for Awareness and Orientation Support in Occluded Multi-Room Settings. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–11, Honolulu HI USA, April 2020. ACM.

[109] Yuhang Zhao, Cynthia L. Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. Enabling People with Visual Impairments to Navigate Virtual Reality with a Haptic and Auditory Cane Simulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14, Montreal QC Canada, April 2018. ACM.

[110] Alexa F. Siu, Mike Sinclair, Robert Kovacs, Eyal Ofek, Christian Holz, and Edward Cutrell. Virtual Reality Without Vision: A Haptic and Auditory White Cane to Navigate Complex Virtual Worlds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, Honolulu HI USA, April 2020. ACM.

[111] Domenick Mifsud, Chris Wickens, Michael Maulbeck, Peter Crane, and Francisco 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, September 2022.

[112] Uwe Gruenefeld, Abdallah El Ali, Susanne Boll, and Wilko Heuten. Beyond Halo and Wedge: visualizing out-of-view objects on head-mounted virtual and augmented reality devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–11, Barcelona Spain, September 2018. ACM.

[113] Hiroto Oshimi, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, and Kiyoshi Kiyokawa. LocatAR: An AR Object Search Assistance System for a Shared Space. In *Augmented Humans Conference*, pages 66–76, Glasgow United Kingdom, March 2023. ACM.

[114] Nicolas Barbotin, James Baumeister, Andrew Cunningham, Thierry Duval, Olivier Grisvard, and Bruce 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)*, pages 213–221, Christchurch, New Zealand, March 2022. IEEE.

[115] Weiquan Lu, Henry Been-Lirn Duh, Steven 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, March 2014.

[116] Christopher D. Wickens, Domenick Mifsud, Richi Rodriguez, and Francisco R. Ortega. Mitigating the Costs of Spatial Transformations With a Situation Awareness Augmented Reality Display: Assistance for the Joint Terminal Attack Controller 3-17. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 65(4):651–662, June 2023.

[117] Camille Truong-Allié, Alexis Paljic, Alexis Roux, and Martin Herbeth. User Behavior Adaptive AR Guidance for Wayfinding and Tasks Completion. *Multimodal Technologies and Interaction*, 5(11):65, October 2021.

[118] Sabine Janssen, Jaap de Ruyter van Steveninck, Hizirwan S. Salim, Helena M. Cockx, Bastiaan R. Bloem, Tjitske Heida, and Richard J. A. van Wezel. The Effects of Augmented Reality Visual Cues on Turning in Place in Parkinson's Disease Patients With Freezing of Gait. *Frontiers in Neurology*, 11:185, March 2020.

[119] Nicklas H. Christensen, Oliver G. Hjermitslev, Frederik Falk, Marco B. Madsen, Frederik H. Østergaard, Martin Kibsgaard, Martin Kraus, Johan Poulsen, and Jane Petersson. Depth cues in augmented reality for training of robot-assisted minimally invasive surgery. In *Proceedings of the 21st International Academic Mindtrek Conference*, pages 120–126, Tampere Finland, September 2017. ACM.

[120] Minh Tien Phan, Indira Thouvenin, and Vincent Fremont. Enhancing the driver awareness of pedestrian using augmented reality cues. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1298–1304, Rio de Janeiro, Brazil, November 2016. IEEE.

[121] Benjamin Volmer, James Baumeister, Raymond Matthews, Linda Grosser, Stewart Von Itzstein, Siobhan Banks, and Bruce H. Thomas. A Comparison of Spatial Augmented Reality Predictive Cues and their Effects on Sleep Deprived Users. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 589–598, Christchurch, New Zealand, March 2022. IEEE.

[122] Benjamin Volmer, James Baumeister, Stewart Von Itzstein, Matthias Schlesewsky, Ina Bornkessel-Schlesewsky, and Bruce H. Thomas. Event Related Brain Responses Reveal the Impact of Spatial Augmented Reality Predictive Cues on Mental Effort. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–18, 2022.

[123] Björn Schwerdtfeger, Rupert Reif, Willibald A. Günthner, and Gudrun Klinker. Pick-by-vision: there is something to pick at the end of the augmented tunnel. *Virtual Reality*, 15(2-3):213–223, June 2011.

[124] Frank Biocca, Arthur Tang, Charles Owen, and Fan Xiao. Attention funnel: omnidirectional 3D cursor for mobile augmented reality platforms. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1115–1122, Montréal Québec Canada, April 2006. ACM.

[125] Robin Hanson, William Falkenström, and Mikael Miettinen. Augmented reality as a means of conveying picking information in kit preparation for mixed-model assembly. *Computers & Industrial Engineering*, 113:570–575, November 2017.

[126] Jonathan Wieland, Rudolf C. Hegemann Garcia, Harald Reiterer, and Tiare Feuchtner. Arrow, Bézier Curve, or Halos? – Comparing 3D Out-of-View Object Visualization Techniques for Handheld Augmented Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 797–806, Singapore, Singapore, October 2022. IEEE.

[127] John Luksas, Kelsey Quinn, Joseph L. Gabbard, Mariam Hasan, Janet He, Neha Surana, Moustafa Tabbarah, and Nishant Kishan Teckchandani. Search and Rescue AR Visualization Environment (SAVE): Designing an AR Application for Use with Search and Rescue Personnel. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 488–492, Christchurch, New Zealand, March 2022. IEEE.

[128] Jad Al Rabbaa, Alexis Morris, and Sowmya Somanath. MRsive: An Augmented Reality Tool for Enhancing Wayfinding and Engagement with Art in Museums. In Constantine Stephanidis, editor, *HCI International 2019 - Posters*, volume 1034, pages 535–542. Springer International Publishing, Cham, 2019. Series Title: Communications in Computer and Information Science.

[129] Stephanie Arevalo Arboleda, Franziska Rücker, Tim Dierks, and Jens Gerken. Assisting Manipulation and Grasping in Robot Teleoperation with Augmented Reality Visual Cues. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, Yokohama Japan, May 2021. ACM.

[130] Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, 1988.

[131] Sandra G Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *HUMAN FACTORS AND ERGONOMICS SOCIETY 50th ANNUAL MEETING*, page 5, 2006.

.1 Demographics Questionnaire

1. Please enter your participant number (this will be given to you by the researcher):

2. Please enter your age:

3. What is your gender?

    - Male
    - Female
    - Non Binary/Third Gender
    - Prefer not to say

4. What is your current academic status

    - Undergrad
    - Recent Grad
    - Master's Student
    - PhD
    - Non student

5. What is your major or profession? (Please type N/A if not applicable).

6. Have you ever used an AR headset?

7. If you answered yes please explain:
8. Have you consumed any of the following in the past 24 hours?
    - Caffeine
    - Alcohol
    - Nicotine
9. On average how many hours a week do you use a computer (personal or work)?
10. Have you ever used a VR headset?
11. If you answered yes which headsets and for how long have you used a VR headset?
12. Have you used a Microsoft Kinect or WiiMote?
13. Indicate the average number of hours per week you spend playing the following types of games:
    - Action-Adventure Games (i.e. Uncharted)
    - Platformers
    - Shooters
    - Fighting games
    - Rhythm Games (i.e. Guitar Hero)
    - Puzzle Games
    - RPGs
    - Simulation Games
    - Strategy Games
    - Sports Games
    - Online Games

- Other please specify with hours

14. Indicate the number of hours per week you spend playing games with the following controllers:

    - Steering Wheel
    - Keyboard and mouse
    - Nintedo DS or 3DS
    - Gameboy or Gameboy color
    - Gameboy advance
    - Touchscreen (i.e. smartphone or tablet)
    - Wii remote
    - WiiU Gamepad
    - N64 Controller
    - Nintendo Gamecube Controller
    - Playstation Controller
    - Playstation Portable (PSP) or Playstation Vita
    - Xbox, Xbox 360, or Xbox One Controller
    - Xbox Kinect
    - Leap Motion 3D controller
    - Oculus Rift
    - Other please specify along with hours

.2 Ranking and NASA TLX

1. Please rank each cue by preference:

- No Cue
- 2D Wedge
- 3D Arrow
- Gaze Line

2. Please rate the perceived task load of the NO CUE condition:
   - How mentally demanding was the task?
   - How physically demanding was the task?
   - How rushed was the pace of the task?
   - How successful were you in accomplishing the task?
   - How hard did you have to work to accomplish your level of performance?
   - How insecure, discouraged, irritated, stressed, or annoyed were you?

3. Please rate the perceived task load of the 3D ARROW condition:
   - How mentally demanding was the task?
   - How physically demanding was the task?
   - How rushed was the pace of the task?
   - How successful were you in accomplishing the task?
   - How hard did you have to work to accomplish your level of performance?
   - How insecure, discouraged, irritated, stressed, or annoyed were you?

4. Please rate the perceived task load of the 2D WEDGE condition:
   - How mentally demanding was the task?
   - How physically demanding was the task?
   - How rushed was the pace of the task?

- How successful were you in accomplishing the task?
- How hard did you have to work to accomplish your level of performance?
- How insecure, discouraged, irritated, stressed, or annoyed were you?

5. Please rate the perceived task load of the GAZE LINE condition:

   - How mentally demanding was the task?
   - How physically demanding was the task?
   - How rushed was the pace of the task?
   - How successful were you in accomplishing the task?
   - How hard did you have to work to accomplish your level of performance?
   - How insecure, discouraged, irritated, stressed, or annoyed were you?