Glanceable AR: Evaluating Information Access Methods for Head-Worn Augmented Reality

Feiyu Lu* Shakiba Davari† Lee Lisle‡ Yuan Li§ Doug A. Bowman¶

Center for Human-Computer Interaction
Department of Computer Science, Virginia Tech, Blacksburg, VA, United States

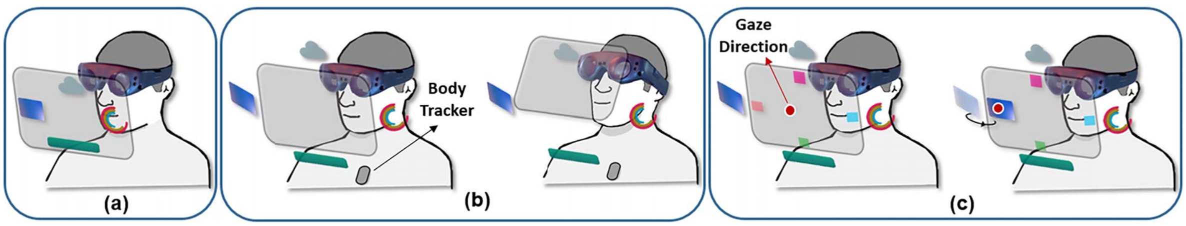

Figure 1: Three Glanceable AR interfaces proposed in this paper: (a): Eye-Glance interface (HUD), in which virtual content resides at the edge of FoV and is fixed to user's head; (b) Head-Glance interface, in which virtual content is invisible in the forward direction, but can be accessed through turning one's head to the periphery; (c): Gaze-Summon interface, in which virtual content can be summoned into FoV by gazing at the periphery for 0.5 second.

Abstract

Augmented reality head-worn displays (AR HWDs) have the potential to assist personal computing and the acquisition of everyday information. In this research, we propose Glanceable AR, an interaction paradigm for accessing information in AR HWDs. In Glanceable AR, secondary information resides at the periphery of vision to stay unobtrusive and can be accessed by a quick glance whenever needed. We propose two novel hands-free interfaces: "head-glance", in which virtual contents are fixed to the user's body and can be accessed by head rotation, and "gaze-summon" in which contents can be "summoned" into central vision by eye-tracked gazing at the periphery. We compared these techniques with a baseline heads-up display (HUD), which we call "eye-glance" interface in two dual-task scenarios. We found that the head-glance and eye-glance interfaces are more preferred and more efficient than the gaze-summon interface for discretionary information access. For a continuous monitoring task, the eye-glance interface was preferred. We discuss the implications of our findings for designing Glanceable AR interfaces in AR HWDs.

Index Terms: Human-centered computing—Mixed / augmented reality; Human-centered computing—User interface design

1 Introduction

People encounter a variety of information needs in their daily lives [7, 9]. Information could be needed on the go to assist in decision-making and execution of certain tasks (e.g., check availability on calendars to arrange a meeting, or check the to-do list to decide activities in a day) [33]. Ideally, we would want such information to be always available and easily accessible but stay out of the way when unneeded to avoid disturbing our tasks at hand.

*e-mail: feiyulu@vt.edu
†e-mail: sdavari@vt.edu
‡e-mail: llisle@vt.edu
§e-mail: yli92@vt.edu
¶e-mail: dbowman@vt.edu

Mobile phones, as the most pervasive personal computing device nowadays, provide convenient ways to obtain information [9, 33]. However, its handheld nature and touch interfaces could introduce limitations in some scenarios [33]. For example, if users want to check today's calendar using a smartphone, they will need to perform the following four steps: (1) pull out the phone from the pocket/purse; (2) unlock the phone; (3) find and open the calendar app, and (4) check for the information they need. If hands are pre-occupied or users have to keep their eyes on other tasks, such procedures could be time-consuming, challenging, or even dangerous. Wearables, especially smartwatches, are capable of displaying information, such as the next calendar event, to users without the need for all the steps above. However, the small screen size limits the amount of information being presented at any time.

Augmented reality head-worn displays (AR HWDs) are becoming more lightweight and powerful. They may eventually replace our smartphones as the primary way to access information. AR HWDs have been long considered as a future way of acquiring information. Back in 2002, Feiner envisioned the future of AR HWDs as "much like telephones and PCs", and displaying information "that we expect to see both at work and at play" [14].

Current AR systems such as the Microsoft HoloLens and Magic Leap One are limited to a single visible application at a time. However, to become versatile general computing devices, AR systems also need to support continuous access to a wide variety of content. All-day AR users will want to check the weather, read their email and social media feeds, and check the calendar without the need to close and open applications each time. We believe this multi-tasking consumption of information will become the primary mode of use of future wearable AR devices. Lages and Bowman proposed an adaptive walking interface for AR HWDs, in which general information is represented in 2D windows. They are able to follow the users around, and adapt themselves to the environment [22]. Their results shed light on how AR HWDs can allow users to bring their personal workspaces anywhere for everyday information access. However, research is still needed on how information can be presented to users in an unobtrusive way, and how we could efficiently and naturally access the information we need in AR HWDs.

In this research, we propose *Glanceable AR*, an interaction paradigm for accessing information in AR HWDs. In Glanceable

2642-5254/20/$31.00 ©2020 IEEE
DOI 10.1109/VR46266.2020.00118

AR, information resides at the periphery to stay unobtrusive, and it can be accessed by a quick glance when needed. We propose two hands-free interfaces within this paradigm. In the *head-glance* interface, virtual content is fixed to the user's body and can be accessed by turning one's head to the periphery. With the *gaze-summon* interface, virtual content can be "summoned" to central vision by eye-tracked gazing at peripheral targets. We evaluated these interfaces, together with a baseline interface *eye-glance*, a heads-up display (HUD) interface, in two dual-task scenarios, in which participants were asked to either answer questions from virtual applications or monitor a virtual basketball scoreboard (secondary task), and at the same time follow a virtual human walking in a large room (primary task).

The contributions of this work include: (1) the Glanceable AR paradigm to acquire secondary information in AR HWDs; (2) three hands-free interfaces proposed under the paradigm; and (3) comparison of the three interfaces in a dual-task walking scenario.

2 Related Work

2.1 Accessing information in AR HWD

DiVerdi et al. proposed ARWin, a desktop workspace augmented by virtual content [10]. The system allows users to manage virtual content and access information such as time, weather report, calendar and web browsers in virtual windows. Valimont et al. compared AR HWDs with interactive videos and printed instructions, and found that AR could lead to more efficient acquisition of information and formation of long-term memory [35]. Lages and Bowman explored how to manage information in an adaptive way while walking [22]. Their results emphasized the importance of customized virtual content positioning, since virtual content could block the real world. Virtual content needs to be presented in a way that does not block central vision, but is easily accessible at the same time. Rhodes proposed the ambient and agent interfaces, in which the amount of information presented to users is managed dynamically based on the cognitive load of the tasks that users are engaging in [3, 31]. For example, if users are engaged in a high-level task, a low-level ambient output is desired. Lindlbauer et al. proposed a real-time approach to determine which information to display, level of detail (LoD) of the information, and where they are displayed based on the user's current cognitive load and knowledge about the tasks and environment [23]. They evaluated their approach in a dual-task scenario, and showed that it reduced the number of interactions to accomplish the secondary task by 36%. In this research, instead of dynamically changing the virtual contents to reduce cognitive effort and avoid occlusion, we aim to explore a new way of accessing information in AR HWDs through glancing at the periphery. In the next two sections, we will elaborate on the nature of peripheral awareness, and why placing content in the periphery could be beneficial.

2.2 Peripheral awareness

In information design, there are multiple strategies for keeping users aware of important information. The peripheral awareness strategy places information into the user's peripheral attention in a way that it is easily accessible without being distracting [6]. Weiser and Brown defined the "periphery" as "what we are attuned to without attending to explicitly" [37, 38]. By placing content in the periphery, users are empowered to take control of the content and initiate interaction based on their needs [25, 37]. For example, users can keep the content in the periphery to make it less obtrusive and important, but bring the content into the central vision to prioritize it and focus on it when needed [37].

In applications requiring situation awareness, peripheral vision is widely seen as effective for non-attentive monitoring of qualitative states or changes in simple secondary information [12]. Embedding information in the periphery can give users opportunities to learn more, do a better job, or keep track of less important tasks [24].

2.3 Accessing information in the periphery

The "Dangling String", created by artist Natalie Jeremijenko, was referred to by Weiser and Brown as an early example of placing information in the periphery for awareness [37]. After Weiser, there has been increasing interest in designing interfaces that engage the periphery of attention. Peripheral displays are displays that are not intended to interrupt user's primary tasks, and allow quick intake of information through a glance at the periphery [26]. InfoCanvas is a peripheral display that allows communication of quantitative information through visual representations of known objects on a secondary display [28]. Sideshow and Scope are peripheral displays that sit at the border of the screen showing general information such as incoming emails and tasks [6, 36]. ShutEye is a peripheral display that encourage awareness of healthy sleep behaviors [2]. Ambient displays are a specific type of peripheral displays that utilize physical materials or lifelike forms distributed in the environment to convey information [17, 18, 40]. These projects emphasized the importance of information awareness, and gauged the viability of obtaining peripheral information in a controlled space.

When it comes to everyday life, information needs could happen on the go in a less-controlled manner. Motivated by this, researchers have explored wearables as an approach to obtain peripheral information. Enriko et al. developed eye-q, an untracked HWD that delivers notifications to wearers by altering brightness and flashing patterns of LEDs in the periphery [8]. The display is lightweight and cost-effective, but is also limited in the amount of information it is able to convey. Drik et al. proposed WatchThru, a wrist-worn peripheral display with AR capabilities to enable better expressivity with a richer set of interactions [39]. However, it requires extra hardware to track the device and the user's head. AR HWDs, due to their integrated tracking and display, have the potential to assist the acquisition of information in the periphery as a wearable device. To our knowledge, few studies have explored this possibility.

3 Glanceable AR

According to Matthews et al., visual displays have the quality of being *Glanceable* if they enable users to understand information quickly with low cognitive effort [26]. Glanceable interfaces allow users to access and understand information displayed to them with a quick glance. Previous research in glanceable display design has mainly focused on how to alter visual features of information (e.g., size, position, contrast, or shape) to make it quickly understandable [25]. For example, a mail app icon may display a badge with the current number of unread messages.

However, designing glanceable interfaces requires careful consideration not only about how to present information visually, but also about how to access the information effectively. Because AR HWDs, unlike current mobile and wearable devices, are worn on the head and always visible, it is challenging to design efficient and effortless interactions to access information with these devices. In this section, we introduce Glanceable AR, a paradigm for designing glanceable interfaces in AR HWDs. We detail the design principles and propose three interfaces within this paradigm.

3.1 Design Principles

Future AR HWDs will likely be always-on wearable displays that can be used on the go. Thus, our vision is to design interfaces for always-on AR HWDs in which virtual content is unobtrusive when not needed, but highly glanceable. To achieve this goal, we adopted four design principles for our Glanceable AR interfaces:

DP1. Virtual content should be fixed to the user.
DP2. Virtual content should be spatially distributed in the periphery.
DP3. The interaction to access virtual content should be hands-free.
DP4. Information should be accessed by glancing quickly at the periphery.

3.1.1 DP1. Fixed to Users

Feiner et al. described three different ways to register windows in AR: (1) *world-fixed*: fixed to locations and objects; (2) *display-fixed*: fixed related to the HWD; and (3) *surround-fixed*: fixed to positions surrounding the user's body [13]. Although *world-fixed* is the most applied setting to manage information in current AR HWDs (e.g., like the window/menu system in Microsoft HoloLens and Magic Leap One), a strict world-fixed layout does not allow mobility, which makes it hard to access information while users are moving or when users might be in many different locations over time. Lages and Bowman suggested that, if AR is to be truly mobile in the future, interfaces should reflect the way we seamlessly move around the world [22]. Therefore, the first design consideration of Glanceable AR is that the virtual content needs to follow the users to ensure availability of the information. This may correspond to either display-fixed or surround-fixed layouts.

3.1.2 DP2. Spatial Distribution

Billinghurst and Starner proposed two metaphors of accessing information in AR HWDs [3]. One is fixed display, in which all information is presented at once in the same form and position, irrespective of which direction the user is looking at. The other one is virtual spatial display, in which information is displayed on a virtual cylinder surrounding the user, and users can either rotate the cylinder or look around to access information they need. The virtual spatial display was proven to be 30% faster to locate information than the fixed display, because users were able to locate the content using their innate spatial memory. Similarly, in Glanceable AR, content is distributed in different directions in the periphery. Our current prototypes make use of four directions: up, down, left and right, and allocate one piece of content to each direction. By doing this, we hope to alleviate the problem of information overload, and make it more efficient for users to locate the contents around them.

3.1.3 DP3. Hands-free Interaction

Handheld controllers, gestures, and speech are the most heavily used input methods for interacting in AR HWDs. Handheld controllers are not very practical for all-day wearable AR, and cannot be used when the hands are needed for other tasks. Hand gesture input requires the hands to be visible by the headset cameras, which could be cumbersome and tiring. Bare-hand interactions have also been proven to pose issues in social acceptance while used in public space [32]. In Glanceable AR, therefore, we suggest that interactions used to access information need to be hands-free. Although voice input is hands-free and has been shown to be effective in some scenarios, it could disturb other people in a shared space. The performance of voice recognition could also be affected by the noise level around the users. Gaze-driven interfaces have been extensively exploited in the context of gameplay, information placement and other types of interactive tasks in both AR and VR (e.g., [11, 27]). In near-eye displays specifically, eye-tracking has been proven to be faster compared to finger-pointing [34]. Since modern AR HWDs (e.g., Microsoft HoloLens2 and Magic Leap One) have embedded advanced sensors to track users' head and eye movement, head-based and gaze-based interactions could be considered as appropriate options for accessing content in AR HWDs.

3.1.4 DP4. Glancing at the Periphery

In Glanceable AR, we define "periphery" as the space at the edges of or beyond the field of view (FoV) of the AR display. Information in the periphery can be easily accessed by glancing. In everyday activities, glancing allows people to quickly obtain information in the periphery [4]. It is unobtrusive in that one can glance without other people noticing, which could be beneficial in protecting privacy and lead to increased social acceptance. In Glanceable AR, content resides at the periphery of attention, and glancing (with head or eye movements) is the primary way of accessing information.

3.2 Interfaces

In this section, we propose three Glanceable AR interfaces: the eye-glance, head-glance and gaze-summon interfaces.

3.2.1 Eye-Glance Interface

The most basic realization of the Glanceable AR approach is a simple Heads-Up Display, which we call an "eye-glance" interface (see Fig. 1a). In an eye-glance interface, contents are placed at the edges of the display's FoV and are fixed to the display. The contents are always visible irrespective of the user's head and body movements, and accessing the information is as simple as moving the eyes to look at it. Eye-glance interfaces are popular in some commercial products (e.g., Google Glass and Focal smart glasses[1]) as a strategy to display information to users. We consider eye-glance to be a baseline interface in that it does not rely on any sensors or tracking, which makes it a versatile and low-cost option. It has the benefits of information being highly visible and accessible, but it may also be obtrusive, occlude the real world, and cause the problem of information overload. Fig. 2a shows an implementation of the eye-glance interface in AR.

3.2.2 Head-Glance Interface

To make content more unobtrusive, the head-glance interface places content outside the user's forward field of view and fixes the content to the user's body rather than the display. To access content, users simply turn their heads towards one side of the periphery (see Fig. 1b). For example, looking up at the sky might allow the user to see information about the weather. As a result of this design, users have a clear view that is not blocked by any virtual content when they are looking forward (relative to the body orientation). However, this interface requires independent tracking of the orientations of both the head and body. Fig. 2b illustrates an AR implementation of the head-glance interface.

3.2.3 Gaze-Summon Interface

The third interface we propose is called the gaze-summon interface. It utilizes the gaze-contingent interaction metaphor, in which the manner of displaying information is presented adaptively based on the user's gaze direction [15, 30]. Based on our review, gaze-contingent interactions are underexplored in AR HWDs. They have the potential to improve the efficiency of obtaining information in future everyday AR displays, when virtual content is likely to be cluttered and overloaded in the physical space surrounding users.

In the gaze-summon interface, content is fixed to the display, but is initially invisible (outside the FoV) to avoid occluding the real-world. To access the content, instead of turning one's head, users need to move their eyes to gaze at the edge of the FoV of the AR HWD. This action "summons" the information, causing it to move into the visible area of the display (see Fig. 1c). We found through informal testing that users have trouble knowing where to look without a visual target. To help users locate the activation areas, small, translucent visual targets are shown. To avoid the effects of eye tracking jitter and the "Midas Touch effect" [20], we use a dwell technique [19]. To summon the content, users are required to dwell their gaze on the target for a short period of time. A shorter dwell time could lead to faster information access, but it could also increase the number of false positives. After iterative testing, we found that a dwell time of 0.5 seconds led to a good balance of speed and accuracy. Fig. 2c illustrates the gaze-summon interface.

While the eye-glance and gaze-summon interfaces might not meet a strict definition of AR, since content is registered to the display

[1] https://www.bynorth.com/focals

Figure 2: AR implementations of: (a) eye-glance interface; (b) head-glance interface, in which user looks up at the sky to acquire information about the weather; (c) gaze-summon interface, in which user gazes at the top visual target to acquire information about the weather; and (d) virtual basketball scoreboard used in our experiment's monitoring task.

Interfaces	Accessibility	Awareness	Unobtrusiveness
Eye-Glance	High	High	Low
Head-Glance	Low	Low	High
Gaze-Summon	Medium	Low	Medium

Table 1: Characteristics of the three interfaces in terms of ease of access, awareness of information, and unobtrusiveness of display.

rather than to the real-world [1], prior AR research has supported the use of a display-fixed frame of reference as an appropriate way to present 2D information in the real-world environment [13, 21].

3.2.4 Tradeoffs Among Interfaces

Table 1 shows our hypotheses about the three interfaces in terms of ease of information access, awareness of information change, and unobtrusiveness of information display. The eye-glance interface makes it easy to access the information with a quick glance at the periphery. It should lead to high awareness of changes in the information by making them always visible, but this same property makes the display obtrusive. The head-glance interface makes contents unobtrusive by placing them in the periphery so they are not visible in the forward direction. This should allow users to be able to pay more attention to things happening in the physical world. However, this comes at the cost of reduced awareness of changes in the virtual content, and information access is also more physically demanding, as users need to turn their heads. The gaze-summon interface is a compromise between the two in terms of unobtrusiveness and accessibility. The visible gaze targets make the content less obtrusive than the eye-glance but more obtrusive than the head-glance interface, while the gaze-based summoning is physically less demanding than head turning but more difficult than the quick glances required by the eye-glance. Like the head-glance interface, the gaze-summon interface lowers the awareness of information change since information is only made visible on demand.

In different contexts of use, these three properties of Glanceable AR interfaces may be valued differently. For example, if the information is non-critical or does not change often, awareness may be less important. When paying attention to the physical world is required, unobtrusiveness may be the most important interface property.

4 Experiment

To explore the benefits and limitations of the interfaces in a particular context of use, we evaluated them in an empirical user study. Participants were asked to perform a primary walking task while doing a secondary task using each of the three interfaces in turn. We aimed to reach a deeper understanding of the tradeoffs among the three techniques in terms of how efficiently participants could access the information they need for the secondary task and how use of the interfaces would affect their performance on the primary task in the physical world.

4.1 Participants

We recruited 18 participants (4 females) between 19 and 23 years old (M = 20.72, SD = 1.23) from our local university. Two of them did not have prior experience with AR before the experiment. Participants all had near-perfect vision with or without contact lenses (we did not allow eyeglasses because they do not work well with the AR headset we used).

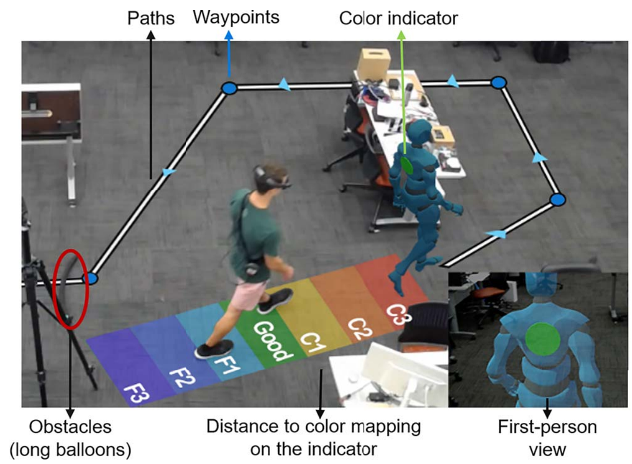

Figure 3: Primary task: follow a virtual human while keeping an ideal distance and avoiding obstacles. Paths and way-points were pre-defined in the program and were invisible to participants. A color indicator on the back of the virtual human showed participant's distance to it. The distance to color mapping is shown on the ground in the image; this was not visible to participants during the experiment.

4.2 Tasks

We used a 3×2 within-subjects design with two independent variables: three *interfaces* (eye-glance, head-glance and gaze-summon), and two *secondary tasks* (discretionary information access and monitoring), yielding six total conditions. Latin square counterbalancing was applied to the order in which interfaces were used.

4.2.1 Primary Task

The primary task was to physically walk in order to follow a virtual human rendered in the AR display. The virtual human walked along pre-defined paths in a 1160 square-foot room (around 108 square meters) for 2 minutes, and changed its speed randomly during the walk. We used spatial sound in the AR headset to represent the virtual human's footsteps in order to help participants locate the virtual human. Participants started each trial directly behind the virtual human. On the back of the virtual human, there was a colored panel indicating the distance between it and the participant. Participants were asked to keep a fixed distance (1.3-1.7 meters) to the virtual human (i.e., to keep the color green). Three long balloons, similar in color to the floor, were placed randomly along the walking path each trial. We instructed participants to avoid these obstacles while walking. Fig. 3 illustrates the task in a top-down view.

4.2.2 Secondary Task

There were two different kinds of secondary tasks.

Discretionary Task: The first type of secondary task was to answer questions from the glanceable information. In this task, there were four pieces of virtual content (weather, activity ring, calendar, and trending news; see Fig. 2a) arranged at the top, left, bottom and right sides of the interface respectively. The weather content showed weather status and temperature. The activity ring showed steps walked and calorie consumption. The calendar showed the next upcoming event. The trending news showed random news headlines. Content was placed at a depth of one meter. For the eye-glance interface, they were placed at the edge of the display FoV so that all of the content was clearly readable. For the head-glance and gaze-summon interfaces, virtual content was placed thirty centimeters beyond the edge of the FoV. For head-glance specifically, participants needed to turn their heads around 16 degrees beyond the periphery to fully read the virtual content. Content had different sizes depending on the amount of information being presented. On average, each piece of content occupied around 10.75 degrees horizontal and 3.43 degrees vertical on the display. Questions were asked verbally by the system; for example, the participant might hear, "What is the temperature now?" With this task, we hoped to simulate the everyday scenario in which information queries are initiated through conversations [33]. The content changed randomly every 10-45 seconds, so participants always had to access the information to be sure to provide the correct answer. Participants heard three questions for each piece of content during the two-minute walk, yielding a total of 12 questions. There was at least a 5-second interval between the end of one question and the start of the next, to give participants enough time to acquire information and give their answers. Participants answered verbally, and were instructed to answer the questions as quickly as they could while still giving priority to the primary walking task.

Monitoring Task: In the monitoring task, participants were asked to monitor a virtual basketball scoreboard (see Fig. 2d), and report as soon as possible when they spotted a "lead change" (i.e., when the team that had been trailing took the lead). The virtual scoreboard always appears at the right side of the interface. It was placed at a depth of one meter and occupied around 10.20 degrees horizontal and 3.43 degrees vertical. During the two-minute walk, the score on the scoreboard changed every 5-15 seconds, leading to a total of 12 score changes, of which eight were lead changes (the sequence of changes was the same for all participants). Unlike the discretionary task, in which participants initiated a query when the program asked them to, in the monitoring task, they needed to pay attention as often as possible to the virtual content, while still prioritizing the primary walking task. To accomplish the monitoring task, participants needed to remember and compare who was leading each time the score changed. Participants reported verbally each time they saw a lead change. For the gaze-summon interface, each translucent visual target occupied around 3.65 degrees horizontal and 2.99 degrees vertical on the display.

4.3 Apparatus

The experiment used a Magic Leap One AR HWD. This device has 1280×960 resolution and a 50-degree diagonal FoV. The headset is connected via a cable to the processing and battery unit, which is worn clipped to the user's belt or pocket.

To ensure that the virtual human walked in the desired pattern, we mapped the room with the headset and defined waypoints on the ground. The waypoints would persist and stay in the exact same physical location every time the application was run, after the headset recognized the space.

The head-glance interface required the users' body to be tracked, so we used the Magic Leap controller for registering body orientation. The controller was placed in a 3D printed case attached to a belt to enhance stability. The belt was secured around the participant's waist to ensure that the controller pointed forward throughout the experiment. The controller is magnetically tracked, so it could be accurately positioned by the headset even if it was occluded by the body. Virtual content rotated around the world vertical axis with the origin at user's head. We synchronized the yaw of the controller with the yaw of virtual content, so content was only visible if the user rotated their head. Before using the head-glance interface, we allowed participants to calibrate the pitch angle of the content to allow comfortable access of information.

For the gaze-summon interface, we used the internal eye-tracking sensors in the Magic Leap headset to track gaze direction. The experimental software was developed via Unity 2019.1.0b10 with the SDK provided by Magic Leap.

4.4 Measures

For all six conditions, we logged participants' head positions, head orientations, eye-tracked gaze positions, and distance to the virtual human ten times per second during the two-minute walk. For the head-glance interface specifically, we recorded participants' body orientation as well. Together with a 3D model of the walking space, we were able to not only measure participants' performance on the primary task, but also reconstruct the walking experience for all participants in a playback system for qualitative analysis.

To measure how well participants performed on the walking task, we used a score function to compute a distance score for each trial [29]. The score function was $F_1 + C_1 + 2(F_2 + C_2) + 4(F_3 + C_3)$, where $F_1/C_1|F_2/C_2|F_3/C_3$ represent time (in seconds) spent in the slightly | moderately | extremely too far/close zones (please refer to Fig. 3 for the detailed coding scheme).

For both discretionary and monitoring tasks, we audio-recorded all the sessions to be able to measure how long it took for participants to answer the question (from the time the question audio finished playing) or report the lead change (from the time the lead change occurred). All the sessions were also video-recorded by a Logitech C930e HD Webcam with 1080p resolution mounted high above the room on the wall, so we were able to combine the virtual playback, the video recording, and the audio recordings for more comprehensive observations of user behaviors.

We used System Usability Scale (SUS) and NASA TLX workload questionnaires to gauge the usability and workload of the interfaces [5, 16]. We also asked participants to rank the interfaces for both secondary tasks, and say what they perceived to be good or bad about the three interfaces in the post-study questionnaire.

4.5 Hypotheses

We tested four hypotheses in the experiment:

H1. With the discretionary secondary task, the head-glance and gaze-summon interfaces will result in superior primary task performance compared to the eye-glance, and will be more preferred, since the user can manage when and where to access the information.

H2. With the discretionary secondary task, the head-glance and eye-glance interfaces will result in better secondary task performance than the gaze-summon interface due to the gaze dwell mechanism.

H3. With the monitoring secondary task, the eye-glance and head-glance interfaces will be more preferred and result in superior performance on the primary task compared to the gaze-summon interface, because continuously accessing the basketball score with the gaze-summon interface will draw attention away from the walking task.

H4. With the monitoring secondary task, the eye-glance and head-glance interfaces will result in better secondary task performance than the gaze-summon interface, because it will be difficult to continuously access the score with the gaze-summon interface.

4.6 Experiment Procedures

The experiment was divided into six phases. In the first phase, participants were welcomed upon arrival, and were asked to read and sign the consent form (the study was approved by the Institutional Review Board of the university). Second, they were asked to fill out a pre-study questionnaire to collect demographic information and prior experience with AR. In the third phase, participants were given a detailed introduction to the experiment background, hardware, three interfaces, and the tasks involved in the study. When participants had no further questions, in the fourth phase, we helped participants to put on the AR HWD, and participants were asked to complete two calibration processes: (1) fitting guide program of the Magic Leap One to determine the ideal size of the forehead-pad and nose-pad; and (2) visual calibration program of the Magic Leap One to ensure proper functioning of eye-tracking. In the fifth phase, participants first completed a training run involving only the primary task without the interfaces. Then they experienced each of the six conditions one by one. Before completing the experimental task in each condition, a training session was provided to get participants familiar with the task, interface, and positioning of virtual content. After this, participants completed one experimental trial with the current condition. After each condition, participants were asked to fill out the SUS questionnaire and the NASA TLX workload questionnaire on a tablet computer. Each condition took about five minutes. After finishing all six conditions, in the sixth phase, participants were asked to fill out a post-study questionnaire, in which we asked them about their preferences and what they thought was good or bad about the interfaces. The entire experiment took about 60 minutes in total. Participants were allowed to take a break anytime in between trials.

5 Results

We conducted a series of analyses to test our hypotheses. We decided not to compare between groups with different secondary tasks because the discretionary and monitoring conditions are very different. As such, we separated the data based on secondary task, and used a one-way repeated-measures ANOVA (RM-ANOVA), with *interface* as the only independent variable for all the analyses. A Greenhouse-Geisser correction was applied for violations of sphericity. For qualitative data gathered from questionnaires and recordings, Wilcoxon signed-rank tests were conducted. We applied Bonferroni corrections for all pair-wise comparisons. We used an α level of 0.05 in all significance tests. In the results figures, pairs that are significantly different are marked with * when $p \leq .05$ and ** when $p \leq .01$. For simplicity, we will use the abbreviations "EG", "GS," and "HG" for "eye-glance interface", "gaze-summon interface" and "head-glance interface" for the rest of the paper.

5.1 Primary Task Performance

Fig. 4 shows the distance-keeping score for the three interfaces with the two secondary tasks. For the discretionary task, EG obtained the lowest distance score (i.e., the best primary task performance) (M=38.04, SD=12.45), followed by GS (M=40.09, SD=13.77) and HG (M=44.68, SD=14.94). However, RM-ANOVA did not find statistical significance in distance score among the three interfaces ($F_{2,34} = 1.523, p = .233$).

Similarly, for the monitoring task, EG obtained the best performance (M=30.78, SD=15.03), followed by GS (M=35.90, SD=11.30) and HG (M=38.43, SD=15.61). Our analysis of the main effect of interface was at the margin of being statistically significant ($F_{2,34} = 3.241, p = .051$). Pairwise analysis found that EG had a significantly better score than HG ($p = .039$). No significance was found for HG-GS ($p = 1.000$), and EG-GS ($p = .384$).

Our other measure of primary task performance was obstacle avoidance. Only two participants hit a balloon in the study. Both of them were from the condition of EG interface with monitoring basketball scoreboard as secondary task.

5.2 Secondary Task Performance

In the discretionary task, we collected 12 (number of questions per condition) $\times 18$ (number of participants)$\times 3$ (number of interfaces)

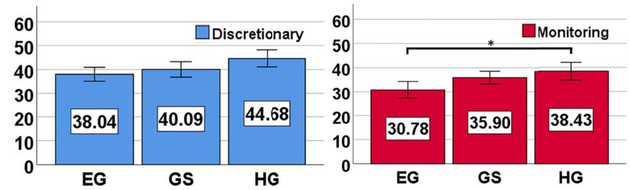

Figure 4: Bar charts of distance score in discretionary (left) and monitoring (right) conditions ($\pm S.E.$)

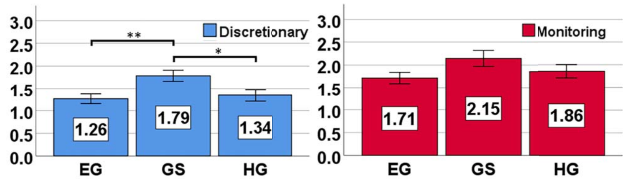

Figure 5: Bar charts of time (in seconds) taken to answer question/report lead changes in discretionary (left) and monitoring (right) conditions. ($\pm S.E.$)

values for the question-answering time measure. We averaged the values for each participant in each condition, and used those averages in our analyses, which leads to a total of 54 data points. Fig. 5 (left) shows the time for each interface. EG resulted in a shorter time (M=1.26s, SD=.45) than both HG (M=1.34s, SD=.57) and GS (M=1.79s, SD=.53). RM-ANOVA revealed a significant effect of *interface* on the time it takes to answer questions ($F_{2,34} = 8.476, p = .001$). Pair-wise comparisons showed that GS led to significantly longer times than both EG ($p = .003$) and HG ($p = .048$). No significant difference was found between EG and HG ($p = 1.000$).

In the monitoring task, none of the participant missed reporting any lead change. As for time taken to report the lead change (Fig. 5 (right)), we again averaged the time for each participant with each interface, leading to a total of 54 data points. EG resulted in a shorter time for reporting lead changes (M=1.71s, SD=.53) than both HG (M=1.86s, SD=.62) and GS (M=2.15s, SD=.75). RM-ANOVA yielded a significant effect of *interface* on the time taken to report lead changes ($F_{2,34} = 3.750, p = .034$). However, pairwise tests did not reveal significant differences among the interfaces (EG-GS: $p = .090$; EG-HG: $p = 1.000$; GS-HG: $p = .210$).

5.3 NASA TLX Workload & SUS Score

Fig. 6 shows a bar chart of the NASA TLX workload sub-scales. Wilcoxon signed-rank tests were conducted to test differences for each subscale. For the discretionary task, EG imposed significantly lower Mental Demand ($Z = -2.731, p = .018$), Effort ($Z = -2.459, p = .042$), and Frustration ($Z = -3.099, p = .006$) as compared to HG. For the monitoring task, EG was rated significantly lower than HG in terms of Physical Demand ($Z = -2.546, p = .033$), Performance ($Z = -2.655, p = .024$) and Effort ($Z = -2.582, p = .03$). EG was also rated significantly lower than GS in terms of Mental Demand ($Z = -2.879, p = .012$) and Effort ($Z = -2.428, p = .045$).

On the SUS questionnaire (Fig. 7), EG (M=88.75, SD=11.61) received a higher score than both GS (M=74.03, SD=17.26) and HG (M=69.58, SD=19.99) in the discretionary condition. A Wilcoxon signed-rank test shows that EG obtained a significantly higher score than GS ($Z = -2.506, p = .036$) and HG ($Z = -2.725, p = .018$). No significant difference was found between GS and HG ($Z = -.785, p = .432$). For the monitoring task, similar results were obtained. Both GS (M=72.50, SD=13.34) and HG (M=74.86, SD=14.21) received lower scores than EG (M=91.39, SD=10.26), and these differences were significant (EG-GS: $Z = -3.182, p = .003$; EG-HG: $Z = -3.007, p = .009$). No significant difference was found between GS and HG ($Z = -.523, p = 1.000$).

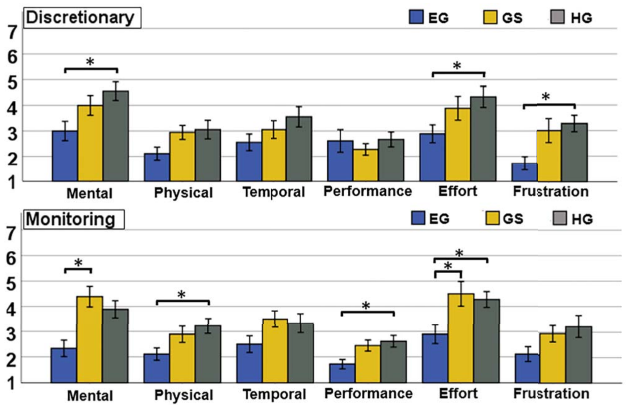

Figure 6: Mean ratings for NASA TLX subscales categorized by interfaces for discretionary (top) and monitoring (bottom) tasks ($\pm S.E.$)

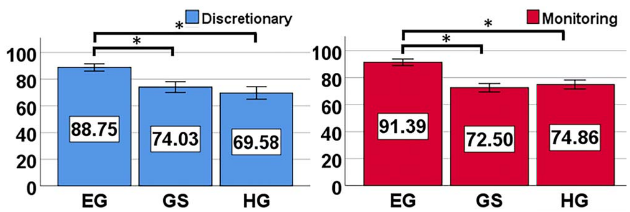

Figure 7: Mean SUS score categorized by interfaces for discretionary (LEFT) and monitoring (RIGHT) tasks ($\pm S.E.$)

5.4 Qualitative Results

5.4.1 Interface Preference

For the discretionary task, participants' choices of preferred interface were somewhat distributed: eight participants (44.44%) preferred EG, six (33.33%) preferred HG, and four (22.22%) preferred GS. On the monitoring task, participants had a clear tendency towards favoring EG. Fifteen participants voted EG (83.33%), two voted HG (11.11%) and one voted GS (5.56%).

5.4.2 Comments on interfaces

When asked what they thought was good or bad about the three interfaces, participants praised EG for being *simple, fast, always there*, and *easy to access*, but also commented that it was *too cluttered, crowded*, and *occluding my view.* For HG, participants liked it because: *[it] has best visibility of the real-world*, *information nearby but not in your face*, and *stuff not in your FoV unless you want it to be*, but disliked that: *[it was] tedious for repeated use*, *[I] have to take my eyes off the walking task*, and *[it was] occluding while turning.* For GS, participants commented *cool to use, futuristic, good visibility of real-world*, and *no need for physical movement in body*, but also *strains my eyes, tracking not always accurate, annoying for repeated use* and *not suitable for monitoring stuff because I have to keep my eyes on the glares [visual targets]*.

When asked whether the number of pieces of virtual information would affect their preferences for the interfaces, fourteen out of eighteen participants (77.78%) gave a positive response. Six of those fourteen (42.86%) commented that more information would make them favor EG interface less. Participants commented that: *lots of windows in the eye-glancing interface would be incredibly annoying* and *more number of windows would mean less available eyesight*. Four participants commented that HG would be a good option for more windows. They commented *lots of windows in the head-glance interface would be easily manageable* and *I wouldn't have to worry about the information in my way unless I wanted it*. Three participants thought GS would work worse when the number of windows increase: *lots of windows in the gaze-summon interface would most likely lead to accidental information pop up* and *there will be plenty of glares [visual targets] within the screen ... it would be kind of hard to tell which boxes [virtual content] would be from which glare [visual targets].*

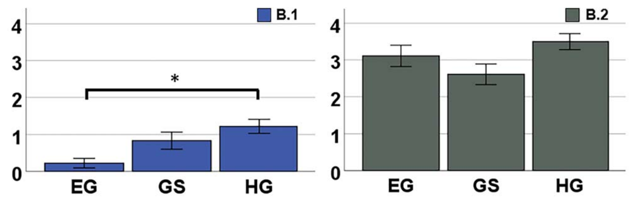

Figure 8: Mean occurrences of (Left) looking around to find virtual human; (Right) adjusting distance with virtual human after a question was asked in discretionary task ($\pm S.E.$)

5.4.3 User Behaviors

We reviewed the video recordings and the reconstructions of the experimental sessions in the playback system to look for interesting or significant behaviors. After an initial review, we decided to count the occurrences of two behaviors within the discretionary task: **B1.** Participants lose track of the virtual human, then look around to find it again (this was likely to happen when a question was asked just before the virtual human made a turn); and **B2.** after a question was asked, participants get too far from or too close to the virtual human, then adjust their distance (likely to happen when a question was asked just before the virtual human changed speed).

Fig. 8 shows the mean number of occurrences for **B.1** and **B.2** categorized by interface. HG (M=1.22, SD=.80) yields more occurrences of **B.1** than GS (M=.83, SD=.98) and EG (M=.22, SD=.55). A Wilcoxon signed-rank test yields a significant difference between EG and HG on **B.1**. (($Z = -3.106, p = .006$)). For **B.2**, HG (M=3.50, SD=.92) and EG (M=3.11, SD=1.23) resulted in more occurrences than GS (M=2.61, SD=1.20). However, no significant difference was found for **B.2**.

6 Discussion

We hypothesized that HG and GS would result in better performance on the primary walking task than EG in the discretionary condition due to better visibility of the virtual human and the real world (**H.1**). Our results did not support **H.1**. No significant difference was found for primary task performance among the three interfaces. We surmise that HG and GS were not advantageous compared to EG due to the characteristics of the primary walking task we chose. To keep an ideal distance, participants needed to pay attention to the back of the virtual human, leading to a primary focus on the forward direction and the central vision. This means that unobtrusiveness was not a major issue because there were no awareness needs in the periphery, so the unobtrusiveness advantages of HG and GS were not reflected in task performance. In contrast, looking away from the forward direction/central vision, as required by HG and, to a lesser extent, GS, could diminish performance on the distance-keeping task. This is also supported by the fact that HG users were more likely to lose track of the virtual human after answering a question, because they needed to look away from the virtual human to acquire information. In addition, all participants obtained relatively low distance scores in the three conditions, and all but two avoided all the obstacles. This indicates that the primary task was relatively easy to perform, even while performing secondary tasks at the same time. Thus, potential disadvantages of the EG interface may not have surfaced during this primary task.

We hypothesized that HG and EG would lead to faster acquisition of information in the discretionary task than GS (**H.2**). Our results supported **H.2** by showing that participants answered questions significantly faster using HG and EG as compared to GS. It is noteworthy that even though EG and HG shared similar secondary task performance, participants perceived that HG posed significantly more Mental Demand, Effort and Frustration than EG. We speculate that this was due to the frequency of information access tasks (12 times in the two-minute walk). With EG, everything is visible directly at the edge of the FoV, and information can be easily accessed through glancing. However, for HG, repeated usage would lead to repeated head and body movements. Due to the unobtrusiveness of information in HG, participants needed to recall the position of content in space each time a question was asked, which creates extra mental workload. Participants commented that HG is not suitable for repeated usage in a short duration, and the repeated head movement led to increased frustration. In real-world scenarios, it is not likely that discretionary acquisition of information would occur so frequently in a short time. If we reduced the number of questions in a trial, we might be able to see more the advantages brought by HG.

Performance of the discretionary task with GS was lower, as we hypothesized. The primary reason for this is that information access is not instantaneous with GS due to the 0.5-second dwell time. Participants also commented that GS is not suitable for frequent uses. Maintaining gaze at the visual target could cause eye strain. In addition, we found that false positives were likely to occur with GS when participants were making a turn while walking. Eye-tracking jitter sometimes increased when users moved their heads quickly, and users also naturally moved their eyes in the direction of the turn before turning their heads and bodies.

Our third hypothesis **H.3** sought to confirm that the advantages of EG and HG could also be seen in a monitoring task, in which repeated (if not continuous) attention is demanded. However, our results only partially supported **H.3**. Most participants preferred EG for the monitoring task, and EG resulted in significantly better walking task performance than HG. EG ensures constant visibility of information, which is a good match for the monitoring task. Also, since we had only one piece of information (the basketball scoreboard), visual cluttering was less likely to be an issue for EG. However, it was somewhat surprising that HG performed poorly on the walking task in the monitoring condition. Although the unobtrusiveness of information in HG might not be ideal for monitoring, we expected that it would give participants the ability to choose when to check the score depending on the status of the primary task. For example, they could choose to check the information quickly when the virtual human was moving at low or constant speed to ensure good performance on the walking task. However, from the playback analysis, we found that despite the fact that we instructed participants to prioritize the walking task, participants tended to continuously check on the score at every opportunity. One participant commented that *I can't see the scoreboard when looking forward [at the virtual human], which makes me feel unsafe*. This constant checking of the score with HG likely led to the significant decrease in walking performance.

Our results failed to support **H.4**. No significant difference was found on time to report lead changes among the interfaces, although GS did take longer than EG and HG in absolute terms. It was expected that GS would not be favored for the monitoring task because continuous gazing is needed on the visual targets to poll the information. We thought that participants might fail to report some of the lead changes while using GS, but none did, which indicates that GS could still be suitable for some monitoring uses.

Overall, it appears that eye-glance was the most optimal interface for the primary and secondary tasks we studied in this experiment, which was surprising to us. Eye-glance interface resulted in the best overall task performance, was rated best for usability and workload, and was preferred by more participants than the other two interfaces. On the one hand, this may indicate that a non-AR wearable display may be all that is needed for secondary information access in some real-world use contexts, especially if the display can be turned off or dimmed when not needed. On the other hand, participants' comments about the interfaces and our own experience lead us to believe that eye-glance is not an ideal solution across a variety of use cases and long-term use. The fixed display on the eye-glance interface can interfere with peripheral awareness of people, objects, and information in the real world, and it can be annoying to constantly have part of one's field of vision occupied with displayed content. The types of Glanceable AR interfaces that require more advanced HWDs, such as gaze-summon and head-glance, are still promising approaches that need further exploration.

7 Limitations & Future Work

There are several limitations to our work. First, we did not vary visual features of the virtual content (e.g., amount of information, 2D/3D, depth, size, color) in our experiment. Future research could explore how different levels of these features could affect user performance on the tasks and user preference for the interfaces. Second, our experiment used an artificial walking task over a short period of time, which does not reflect long-term realistic use. Future longitudinal studies in real-world settings could be used to measure long-term effects of our designs. Third, the AR HWD used in our study, although representing the current state-of-the-art, still has limited FoV and eye-tracking capabilities. Results might be different for future wearable AR glasses. Fourth, we only tested our interfaces with walking as a primary task. Walking is a common everyday behavior, but is relatively simple in terms of workload. In the future, we are planning to test our interfaces with other primary tasks that differ in physical and mental workloads.

We are currently working to improve the head-glance and gaze-summon interfaces based on the feedback we gathered from participants. Our primary goal is to increase unobtrusiveness of Glanceable AR, while still providing efficient information acquisition with minimal effort and enhancing awareness. We plan a follow-up experiment in which we will again compare Glanceable AR interfaces under a more demanding primary task (e.g., driving in a simulator). The primary task will be demanding not only in the central vision, but also in the periphery, so we can evaluate how eye-glance, gaze-summon, and head-glance interfaces perform under such conditions.

8 Conclusions

In this research, we proposed Glanceable AR, an information access paradigm for AR HWDs. We proposed two novel hands-free interfaces using head rotation or eye-tracked gaze to access information. We evaluated them in two dual-task scenarios along with a baseline HUD technique. We found that the head-glance and eye-glance interfaces could lead to faster acquisition of information, but that the head-glance and gaze-summon interfaces are not suitable for frequent usage over a short duration, which could lead to increases in physical and mental workload. The eye-glance interface was more preferred when continuous attention is required on the content being displayed.

We believe that AR HWDs will become an important personal computing device to assist our acquisition of everyday information in the near future. Instead of opening a single app at a time and fix contents in space, future AR HWDs should be capable of rapid information access on the go. Design challenges still exist on how to make information acquisition in AR HWDs effortless and efficient without disturbing the tasks we are doing in the real world. Glanceable AR is an important step towards designing easy-to-use interactions to tackle information needs in AR HWDs.

Acknowledgments

We would like to thank all the participants for their time and the reviewers for their comments and suggestions that helped improve our paper. This work has been supported by the Office of Naval Research.

REFERENCES

[1] R. T. Azuma. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385, 1997.

[2] J. S. Bauer, S. Consolvo, B. Greenstein, J. Schooler, E. Wu, N. F. Watson, and J. Kientz. Shuteye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1401–1410. ACM, 2012.

[3] M. Billinghurst and T. Starner. Wearable devices: new ways to manage information. *Computer*, 32(1):57–64, 1999.

[4] J. Birnholtz, L. Reynolds, E. Luxenberg, C. Gutwin, and M. Mustafa. Awareness beyond the desktop: exploring attention and distraction with a projected peripheral-vision display. In *Proceedings of Graphics Interface 2010*, pp. 55–62. Canadian Information Processing Society, 2010.

[5] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[6] J. Cadiz, G. Venolia, G. Jancke, and A. Gupta. Sideshow: Providing peripheral awareness of important information. Technical Report MSR-TR-2001-83, September 2001.

[7] K. Church, M. Cherubini, and N. Oliver. A large-scale study of daily information needs captured in situ. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(2):10, 2014.

[8] E. Costanza, S. A. Inverso, E. Pavlov, R. Allen, and P. Maes. Eye-q: Eyeglass peripheral display for subtle intimate notifications. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pp. 211–218. ACM, 2006.

[9] D. Dearman, M. Kellar, and K. N. Truong. An examination of daily information needs and sharing opportunities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pp. 679–688. ACM, 2008.

[10] S. Di Verdi, D. Nurmi, and T. Hollerer. Arwin-a desktop augmented reality window manager. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, pp. 298–299. IEEE, 2003.

[11] A. T. Duchowski. Gaze-based interaction: A 30 year retrospective. *Computers & Graphics*, 73:59–69, 2018.

[12] M. R. Endsley. Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society annual meeting*, vol. 32, pp. 97–101. SAGE Publications Sage CA: Los Angeles, CA, 1988.

[13] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the world: 2d windows for 3d augmented reality. In *ACM Symposium on User Interface Software and Technology*, pp. 145–155, 1993.

[14] S. K. Feiner. Augmented reality: A new way of seeing. *Scientific American*, 286(4):48–55, 2002.

[15] J. H. Goldberg and J. C. Schryver. Eye-gaze-contingent control of the computer interface: Methodology and example for zoom detection. *Behavior research methods, instruments, & computers*, 27(3):338–350, 1995.

[16] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, 2006.

[17] J. M. Heiner, S. E. Hudson, and K. Tanaka. The information percolator: ambient information display in a decorative object. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*, pp. 141–148. ACM, 1999.

[18] D. Holstius, J. Kembel, A. Hurst, P.-H. Wan, and J. Forlizzi. Infotropism: living and robotic plants as interactive displays. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pp. 215–221. ACM, 2004.

[19] R. J. Jacob. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 11–18. ACM, 1990.

[20] R. J. Jacob. Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in human-computer interaction*, 4:151–190, 1993.

[21] R. Jose, G. A. Lee, and M. Billinghurst. A comparative study of simulated augmented reality displays for vehicle navigation. In *Proceedings of the 28th Australian conference on computer-human interaction*, pp. 40–48, 2016.

[22] W. Lages and D. Bowman. Adjustable adaptation for spatial augmented reality workspaces. In *Symposium on Spatial User Interaction*, p. 20. ACM, 2019.

[23] D. Lindlbauer, A. M. Feit, and O. Hilliges. Context-aware online adaptation of mixed reality interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 147–160. ACM, 2019.

[24] P. P. Maglio and C. S. Campbell. Tradeoffs in displaying peripheral information. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 241–248. ACM, 2000.

[25] T. Matthews, J. Forlizzi, and S. Rohrbach. Designing glanceable peripheral displays. *UC Berkeley. Retrieved September*, 23:2015, 2006.

[26] T. L. Matthews, J. Mankoff, and J. Canny. *Designing and evaluating glanceable peripheral displays*. University of California, Berkeley, 2007.

[27] A. McNamara and C. Kabeerdoss. Mobile augmented reality: Placing labels based on gaze position. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 36–37. IEEE, 2016.

[28] T. Miller and J. Stasko. The infocanvas: information conveyance through personalized, expressive art. In *Conference on Human Factors in Computing Systems: CHI'01 extended abstracts on Human factors in computing systems*, vol. 31, pp. 305–306, 2001.

[29] M. Nabioyuni and D. A. Bowman. An evaluation of the effects of hyper-natural components of interaction fidelity on locomotion performance in virtual reality. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments*, pp. 167–174. Eurographics Association, 2015.

[30] S. M. Reder. On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behavior Research Methods & Instrumentation*, 5(2):218–228, 1973.

[31] B. Rhodes. Wimp interface considered fatal. In *IEEE VRAIS'98: Workshop on Interfaces for Wearable Computers*, 1998.

[32] J. Rico and S. Brewster. Gesture and voice prototyping for early evaluations of social acceptability in multimodal interfaces. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, p. 16. ACM, 2010.

[33] T. Sohn, K. A. Li, W. G. Griswold, and J. D. Hollan. A diary study of mobile information needs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 433–442. ACM, 2008.

[34] V. Tanriverdi and R. J. Jacob. Interacting with eye movements in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 265–272, 2000.

[35] R. B. Valimont, D. A. Vincenzi, S. N. Gangadharan, and A. Majoros. The effectiveness of augmented reality as a facilitator of information acquisition. In *Proceedings. The 21st Digital Avionics Systems Conference*, vol. 2, pp. 7C5–7C5. IEEE, 2002.

[36] M. Van Dantzich, D. Robbins, E. Horvitz, and M. Czerwinski. Scope: Providing awareness of multiple notifications at a glance. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 267–281. ACM, 2002.

[37] M. Weiser and J. S. Brown. Designing calm technology. *PowerGrid Journal*, 1(1):75–85, 1996.

[38] M. Weiser and J. S. Brown. The coming age of calm technology. In *Beyond calculation*, pp. 75–85. Springer, 1997.

[39] D. Wenig, J. Schöning, A. Olwal, M. Oben, and R. Malaka. Watchthru: Expanding smartwatch displays with mid-air visuals and wrist-worn augmented reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 716–721. ACM, 2017.

[40] C. Wisneski, H. Ishii, A. Dahley, M. Gorbet, S. Brave, B. Ullmer, and P. Yarin. Ambient displays: Turning architectural space into an interface between people and digital information. In *International Workshop on Cooperative Buildings*, pp. 22–32. Springer, 1998.