

# See or Hear? Exploring the Effect of Visual/Audio Hints and Gaze-assisted Instant Post-task Feedback for Visual Search Tasks in AR

Yuchong Zhang \*

Chalmers University of Technology, Sweden

Adam Nowak †

Lodz University of Technology, Poland

Yueming Xuan ‡

Chalmers University of Technology, Sweden

Andrzej Romanowski §

Lodz University of Technology, Poland

Morten Fjeld ¶

Chalmers University of Technology, Sweden

University of Bergen, Norway



Figure 1: Case study: Visual book-searching task with the aid of AR using HMD. *a)*: Our AR app gives a book title as a search stimulus. *b)*: Instant post-task feedback is provided by a dot smoothly following the eye trajectory. (Note: the dot is bright and highly visible in AR but is dim and difficult to see in print; here, it is located between the first and second shelves of the bookcase). *c)*: The visual hint, as a bright purple arrow, supports the task. The timer is designed for measuring the task time (at the top of this figure). Both the gaze playback and visual hints are rendered in real world space for precise displacement. Please check our supplementary video for the demonstration of the book-searching task.

## ABSTRACT

Augmented reality (AR) is emerging in visual search tasks for increasingly immersive interactions with virtual objects. We propose an AR approach providing visual and audio hints along with gaze-assisted instant post-task feedback for search tasks based on mobile head-mounted display (HMD). The target case was a book-searching task, in which we aimed to explore the effect of the hints together with the task feedback with two hypotheses. **H1**: Since visual and audio hints can positively affect AR search tasks, the combination outperforms the individuals. **H2**: The gaze-assisted instant post-task feedback can positively affect AR search tasks. The proof-of-concept was demonstrated by an AR app in HMD and a comprehensive user study ( $n=96$ ) consisting of two sub-studies, Study I ( $n=48$ ) without task feedback and Study II ( $n=48$ ) with task feedback. Following quantitative and qualitative analysis, our results partially verified **H1** and completely verified **H2**, enabling us to conclude that the synthesis of visual and audio hints conditionally improves the AR visual search task efficiency when coupled with task feedback.

**Keywords:** Augmented reality, Search task, Visual hint, Audio hint, Gaze assistance, Instant post-task feedback, User study

\*yuchong@chalmers.se; yuchongz@kth.se

†203151@edu.p.lodz.pl

‡echoxuan98@gmail.com

§androm@kis.p.lodz.pl

¶fjeld@chalmers.se; Morten.Fjeld@uib.no

## 1 INTRODUCTION

Even as Augmented Reality (AR) continues to mature technically, it has already become a cutting-edge technique for ordinary use in everyday life [34, 63, 65]. AR features the capacity to superimpose virtual context-related information onto the physical world in the form of graphics [35, 36, 61, 64]. This technology is endorsed and applied in numerous fields due to its capability of providing real-time interaction, as well as generating interactive interfaces of visualized digital content [17, 27, 62]. An emerging trend of AR is to deploy this technique to generate capability for conducting specific tasks. The visual search task in AR is a widespread research area which has received considerable attention, while the representations of it can be diverse, including visual element searching [14], real world object searching [10], or general searching [38, 45]. As a means to make the search process more accessible, it is even possible that HMD devices with the most advanced AR technology could soon become as ubiquitous as smartphones [40] to empower the search process to become more accessible. We investigated the specialized case study of a book-searching task in AR, which includes all the essential components of a visual search procedure: visual environment, a particular object (the target book), and distractors (irrelevant books) [48], and is easy to realize.

When performing particular tasks in the context of AR, hints can improve the ultimate task performance. The type of hint, acting as a central and core tool, can vary. A widely-used tool is the visual hint [52], which is a mixture of graphical representations in AR with user interface (UI) actions associated with the physical world [54]. This effectively provides spatial and temporal guidance in AR. Another well-adopted hint is the audio hint, which presented as voice commands or directives can swiftly and clearly aid the user. A number of researchers have exploited audio hints for the purpose of informing, guiding, and directing within their designs [22, 42].

However, even though there is a substantial body of research into how visual and audio hints work in AR tasks, little has focused on search tasks other than exploring the combination of visual and audio hints in one AR framework. In our study, we observe both the separate effect of these two types of hints and exclusively their combined effects in AR searching. The context of a visual book-searching task is engaged in an AR approach with supportive visual or/and audio hints.

Gaze assistance, in particular, has a promising role in AR applications owing to its easy, natural, and fast way of involving people in virtual environments [30]. Gaze-assisted techniques and implementations have been demonstrated to efficiently support AR applications by providing better user experiences (UX) [16] and more accurate target selection techniques [15]. Human eyes are easy to track, and gaze implicitly conveys what people are interested in. Thus, eye-tracking approaches are becoming more pervasive in interactive AR devices, which are mainly mobile head-mounted or hand-held displays (HMDs and HHDs) [59], such as Microsoft HoloLens or Magic Leap [35]. Human gaze is utilized for adjusting and adapting virtual information that is being projected onto the real world. The usage of gaze has been extensively demonstrated to be efficient in visual search tasks. Noteworthy, eye gaze recordings have been harnessed in many AR-led or AR-guided contexts [28, 44, 56]. However, very little research has investigated the combination of visual/audio hints and human gaze within visual search tasks within the same context.

Another imperative aspect of AR is task feedback. At present, there is almost no in-depth research into involving gaze assistance for task feedback. It is crucial for domain users to obtain better task performance in many situations related to AR [9, 46]. According to Vieira et al. [51], task feedback has the potential to provide more specific information needed by the users in AR, which encouraged us to explore its potential influence in the same context. Here, we explore the effect of task feedback through gaze-assistance given that both of these have proven their efficacy in AR while there is an evident gap on the combination of them. Instead of real-time feedback [32], we adopted post feedback [43, 66] in our visual book searching context since it allows people to see their task reflection after a complete searching process [2]. The task feedback in our study is embodied as the playback of the human eye gaze recorded during the book-searching process, which is symbolized as an instant evaluation method. A vision-friendly colored dot smoothly following the path is used to display the eye trajectory (Figure 1.b). Specifically, the gaze playback component enabled by eye-tracking serves as the instant post-task feedback in our study design. It is noteworthy that no relevant research, to our knowledge, has studied the effect of associating visual/audio hints with instant post-task feedback in the same context of AR visual searching.

By realizing these research gaps (combining visual and audio hints; task feedback and gaze assistance; visual/audio hints and task feedback), this paper proposes a complete AR approach for visual search tasks, examines the independent and the combined impact of visual and audio hints, and studies instant post-task feedback through gaze assistance. Under the premises that visual/audio hints yield beneficial effects [3, 8, 23], so as task feedback [9, 58], through the AR book-searching task, we intend to test the hypotheses:

- **H1:** Since visual and audio hints have a positive effect in facilitating AR visual searching performance and decreasing perceived workload in AR, the combination of these two hints has a greater effect than either does individually.
- **H2:** The gaze-assisted instant post-task feedback has a positive effect for task performance and perceived workload reduction in AR visual search tasks.

To examine **H1** and **H2**, we designed a comprehensive between-subject [57] user study with a within-subject factor [7] followed

by statistical analyses with the aid of a specialized-for-searching bookcase. The paper contributes:

- Proposing an AR approach supporting visual and audio hints, as well as gaze-assisted instant post-task feedback for visual search tasks based on mobile HMD;
- Exploring the effect of combining visual and audio hints in contextual AR visual searching;
- Exploring the effect of instant post-task feedback through gaze assistance in the same context.
- Exploring the effect of combining hints and instant post-task feedback in a same AR context.

This paper is organized as follows. Section 2 describes related work of AR with gaze assistance, visual/audio hints, and task feedback. In Section 3, we introduce the details of our proposed AR approach. We also incorporate the particular use case of the visual book-searching task. The entire user study is presented in Section 4, with two sub-studies conducted. Section 5 contains the results from the study and quantitative data analysis along with the qualitative summary. Discussion with limitations follows in Section 6, and conclusions and future work are elaborated in Section 7.

## 2 RELATED WORK

### 2.1 Visual Search Tasks in AR

Visual search tasks are becoming more common in AR because of the capability of projecting additional virtual information on the physical environment. Contreras et al. [10] presented a mobile application with AR encapsulated to admit users to search for desired places, people or events on a university campus. The superiority of AR lay in the fact that it offered certain visual elements that helped users to better locate the required result. Rafiq and colleagues [37] proposed a dynamic AR framework to support an online book-searching task by using mobile augmented data. This framework also introduced a security layer which ensured the protection of sensitive cloud data. Gebhardt et al. [12] utilized gaze movement data to observe the MR object's label in a visual searching process through a reinforcement learning method. Trepkowski et al. [49] presented a series of simulating experiments to investigate how visual search performance is affected by the field of view and information density in AR, indicating that a significant effect was caused by these two factors. Van Dam et al. [50] studied the cues in a drone-based AR signal detection task but found no significant differences across AR cue types. Nevertheless, the effects of hints and task feedback continues to be more valued in AR visual searching.

### 2.2 Hints in AR

Numerous researchers have endeavored to bring different modalities of hints into AR/XR systems as auxiliary tools for reaching desired results. Of these, visual hints [4, 26, 54, 60, 67, 68] and audio hints [18, 22, 47, 68] are the two most common representations used. Arboleda et al. [3] presented augmented visual hints in a robot-involved AR system which aimed to enhance the visual space of the robot operator about the position of the robot gripper in the workspace, where the visual hints were used to improve distance perception and then the manipulation and grasping task performance. For tangible AR, White et al. [54] examined visual hints to enable discovery, learning, and completion of gestural interaction in a tangible environment. Seven visual hint types were generated: text, diagram, ghost, animation, ghost+animation, ghost+text, and ghost+text+animation. Two decades ago, Sawhney et al. [42] harnessed audio information to keep users of wearable devices updated with incoming messages and events. Lyons et al. [22] developed an AR game system named "Guided by Voices", which equipped the user with a narrative sound clip that indicated the scenarios encountered and the correct steps to be taken for proceeding. Interestingly

and more recently, Mulloni et al. [24] found that AR can be integrated with a mobile navigation system, while audio hints can be advisable for eye-free usage. Cidota et al. [8] compared the effects of automatic visual and audio notifications regarding workspace awareness in AR remote collaborating. However, the integration of visual and audio hints in one AR context is still insufficient.

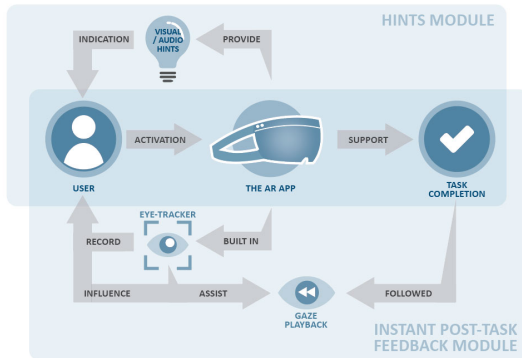


Figure 2: Block diagram: The proposed AR approach with the two core inter-correlated modules: hints (top) and instant post-task feedback (bottom). The AR app builds on the intersection.

### 2.3 AR with Gaze Assistance

Some research has already pioneered gaze-assisted UI to access more contextual information [1, 21, 31, 41]. It has been shown that the optical see-through (OST) HMD which harnesses human gaze with eye-tracking as the interaction metaphor can contribute to efficient results [20]. Interestingly, there are some studies exploring eye-tracking to present menus to aircraft pilots, and adjust their contents based on what the pilot is looking at [33]. In 2005, Curatu et al. [11] proposed a novel conceptualized system adding eye-tracking capabilities to a Head-Mounted Projection Display (HMPD), which was satisfyingly performed from a low-level optical configuration. Three years later, Park et al. [30] pioneered a system which includes a Wearable Augmented Reality System (WARS) to examine their proposition on an experiment in which they selected desirable items in an AR gallery with content mobility. Rivu et al. [40] successfully demonstrated the superiority of eye-tracking, showing that users are more positive in concentrating on their ongoing conversations when there is more gaze interactivity under AR environment. All the relevant research proved that the gaze assistance in AR activates improved perception of people.

### 2.4 Task Feedback in AR

Task feedback evaluation in AR has received compelling attention in the past years. A recent study carried out by Liu et al. [19] evaluated the influence of real-time task feedback in mobile handheld AR, which suggested that significant benefits will emerge if task feedback is engaged during the AR task. Zahiri et al. [58] studied the feedback in AR for supporting a surgical training, finding that most of the users preferred receiving the feedback while the task was being performed. Murakami et al. [25] found that haptic feedback as task evaluation can help users perform more effectively in a wearable AR system for virtual assembly tasks. Clemente et al. [9] demonstrated that continuous visual AR feedback can deliver effective information for the users in their sensorimotor control with a robotic hand, which has implications for amputee assistance in a clinical scenario. Anderson et. al designed a system called YouMove which used interactive At mirrors to train users to record and learn physical movement sequences, where they employed post-stage feedback in their user study. Nonetheless, research regarding post feedback in AR is still

scarce. Even though some researchers have identified the effects of instant post-task feedback in gaze-interaction embodied virtual reality (VR) [39], the gap of linking it with gaze assistance in AR is still obvious, as is the gap of combining it with visual/audio hints. Our study intended to fill the aforementioned gaps and open up future research directions regarding AR visual search tasks.

## 3 SYSTEM DESIGN

### 3.1 Outline of the Proposed AR Approach

The proposed system, featuring an AR app, is an interactive tool for users carrying out a book-searching task. The equipment kit consists of a mobile HMD with built-in eye tracker which is responsible for recording the user's gaze during the task being conducted. An overall block diagram of our proposed system is shown in Figure 2. There are two inter-correlated modules: the hints module and the instant post-task feedback module. The first module was designed to address **H1**, while the second targeted **H2**. The components "User", "The AR app", and "Task completion" are commonly owned by the two modules. The AR app is the core of the system which bears the necessary information needed for the book-searching task including hints and instant post-task feedback. The hints module features a "Visual/audio hints" component while "Eye-tracker" and "Gaze playback" components characterize another module. During the task period, the user's eye motion is entirely captured by the eye tracker and simultaneously recorded. These recordings then act as the feedback for later self-reflection. After the completion of an assigned task, users can obtain data on the total time elapsed and then review the gaze playback to monitor their performance and make improvements for the next task. The central advantages of this system are the simultaneous visual/audio hints designed to enhance task performance, and the instant post-task feedback pathway which offers users the opportunity to make immediate improvements.

### 3.2 Book-searching Task

The book-searching task was implemented with the aid of the proposed AR approach. The AR app encoded a number of book titles which were used as the stimulus in the study. The user with an HMD with a real-time built in eye-tracker would stand in front of the bookcase, to find a specific but randomly determined book (Figure 1.a) after activating the app. In the beginning, users allocated in distinct sub-studies were assigned into particular groups. Depending on the settings, different types of hints with/without instant post-task feedback were then provided. There were two identical book-searching tasks in this study and the gaze playback served as the feedback which was inserted between the two tasks. To precisely control the experimental variables while differentiating the task content, we adopted two congruent bookcases but with totally different books being cleared placed. (Figure 3). The books used for the two tasks are also distinct. The spatial layout of books remained unchanged and unmoved in both bookcases throughout the entire study.

### 3.3 Hints and Instant Post-task Feedback

According to Arboleda et al. [3], visual hints can improve depth perception, which allows people to cognitively better understand their spatial environment. In Wolf et al.'s work [55], animated arrows produced by an AR HMD were used as visual hints to assist dementia patients with navigation. We also chose arrows as the indicators, but showed them in an easily-recognizable color and in a static state. Since all virtual artifacts are displayed in the world space, users' movements do not affect an object's position. As shown in Figure 1.c, the bright-purple arrow floating in the middle of a shelf visually helps the user to locate the target book. This visual arrow is designed to appear within the users' field of view, easily observed by users wearing the HMD. The arrow for the first searching task was positioned on the left side while it appeared on the right side, since the first task started from the left bookcase. Furthermore, the

audio hint also plays a decisive role in guiding users during the searching process. Marquardt et al. [23] demonstrated that AR users can be guided with higher accuracy and within a shorter searching time by incorporating auditory hints. In our developed app, users receive an audio instruction about the approximate location of the targeted book. The auditory information is presented as a clear and simple instructive voice message, allowing the users to finish the task with higher efficiency. This message was merely given once at the beginning of the task to serve as the means of guiding, and not repeated during the latter searching process to avoid being as the external noise disturbance.



Figure 3: User study: An example scene of one participant searching for a book. a): During searching; b): Book found. The two congruent bookcases are used for the two searching tasks.

The functionality and advantages of task feedback have been described in Cao et al.’s findings [6], that furnishing feedback in AR environments offers users more awareness of surrounding information, which ensures task efficiency and correctness. Appropriate feedback in a visually distinguishable form can alleviate their workload and result in positive improvements in task performance [29]. In our study, the instant post-task feedback was provided as a visible playback of gaze recording. As shown in Figure 1.b, the user’s gaze was denoted as a brightly colored dot following the trajectory. During the book-searching phase, gaze trajectories were marked and recorded. After one search task, users then watched the playback of their trajectories, after which they proceeded to the next task.

### 3.4 Apparatus

We used Microsoft HoloLens 2 glasses and the dedicated AR app we developed. The mobile headset weighs 566 grams and has a built-in battery. The see-through holographic lenses, SoC Qualcomm Snapdragon 850, are a second-generation custom-built holographic processing unit with 4 GB RAM memory and 64 GB storage. The users use the glasses freely without any connection to a power supply or external computing device. Built-in eye-trackers allow us precise recording of users’ gaze trajectories. The app was developed with use of Unity3D 2020.3.20f1 and Mixed Reality Toolkit (MRTK).

## 4 USER STUDY

We designed and implemented a comprehensive comparative user study based on the between-subject with coupling within-subject factors. The independent variables were the tasks (within-subject) and the groups (between-subject). The study was comprised of two sub-studies (Study I and Study II) where the participants from Study II received the instant post-task feedback while those from Study I did not. Both studies I and II followed the same between-/within-subject design. Before starting the experiment, we conducted a pre-testing session where several people were invited to test the desired functionalities of the AR app. Some minor adjustments were then made to improve the visual and locational clarity, including the visual hint made into a more vision-friendly arrow, and our testing environment moved to a more spacious and bright function room.

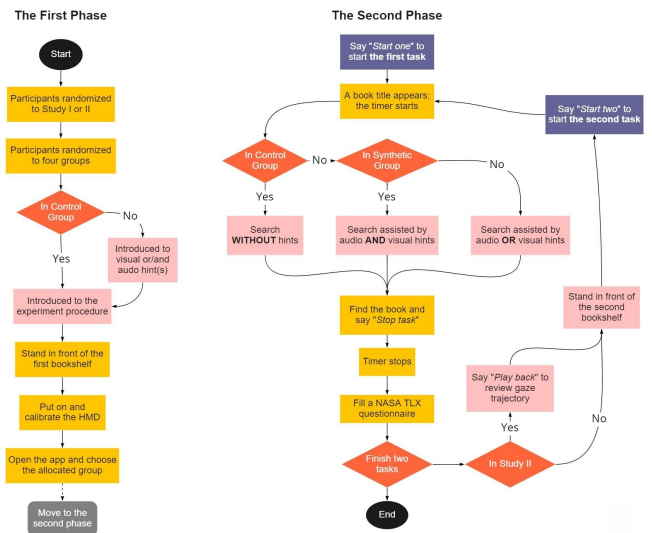


Figure 4: User study: The flowchart of the task procedure. The two phases show how the participants were engaged in the study. The first phase presents the prerequisite of the study while the second phase illustrates the formal searching tasks.

### 4.1 Participants

We recruited 96 participants (different from the people involved in the pre-testing session), of whom most were from a local university. Their ages ranged from 19 to 57 ( $mean = 27.92$ ,  $SD = 7.70$ ); 54 voluntarily self-reported male and 42 self-reported female. There were 48 participants in both Study I (24 males and 24 females, aged from 20 to 44 ( $mean = 26.28$ ,  $SD = 4.64$ )) and Study II (30 males and 18 females, aged from 19 to 57 ( $mean = 29.11$ ,  $SD = 9.20$ )) respectively. No participants decided to quit during the experiment, reported any forms of color blindness, or showed discomfort or rejection of the HMD used in this study. Each participant carried out the experiment independently and correctly. It is therefore reasonable to base our analysis and arguments on the data obtained. The whole study (including I and II) was executed in a bright and spacious function room without any external distractions other than the two bookcases used. There were no additional noises affecting the participants during the whole experimenting process. The acoustic conditions of the function were satisfactory to ensure everyone clearly receiving the audio hint. Each participant was given a small gift for helping with the study afterwards. All of the data collection corresponded to the Covid-19 and ethical rules of the authors’ home universities.

### 4.2 Experiment Design

The participants ( $n=96$ ) were randomly sorted into Study I and Study II ( $n=48$  in each) and again randomly and evenly distributed in four groups in both studies: control, audio, visual, and the combined groups ( $n=12$  in each). The only variation among the groups was the hint difference. We utilized identical settings in the two sub-studies; however, all the groups from Study II received the instant post-task feedback while those from Study I received no such feedback. All participants ( $n=96$ ) executed two book-searching tasks. For the control group, we did not offer any hints. The visual and audio groups received visual (the directive bright-purple arrow pointing out a specific bookcase shelf) or audio (the instructive and articulate voice message (“Please look at the Xth shelf of the bookcase from the top”)) hints respectively. Finally, the combined group was given both sets of hints when performing the book-searching. The AR app was developed by incorporating the precise calculation of the inherent parameters of the bookcase (for example, the height of each

shelf). The height of the participants did not affect the results since the visual arrows were continuously pointing out the correct shelf. The group allocation was employed for controlling the variables to verify **H1** and the hints module in the block diagram (Figure 2). Each group in Study II underwent the feedback phase, where all participants were requested to watch their own gaze recordings from their first book-searching task before undertaking the second task. This was used for identifying the exact role of the instant post-task feedback. The results were to address **H2**, as well as the usefulness of the instant post-task feedback module 2. Everyone was requested to search for different books in two different bookcases in the respective two tasks (Figure 3).

### 4.3 Task Procedure

The duration time for the study of every participant lasted from 10 to 15 minutes. A short and concise pre-training session of book-searching was employed to get all the participants familiarised with the entire environment to maximally eliminate the potential accumulative learning effect among the tasks. The task completion time (TCT), which is defined as the duration from the participant initiates the searching until the book has been found, was measured regarding each task. The complete procedure for our study (including studies I and II), as seen in Figure 4, consisted of two phases:

- The **first phase** started with the introduction of the study and calibration of the HMD. During the introduction, each participant was randomly allocated to one of the four groups followed by an explanatory overview of the experiment procedure by conductor of the study. If participants were not assigned to the control group, the roles of the audio or/and visual hint(s) were introduced and explained. Participants were then told to stand in front of the first test bookcase at a distance of one metre (a marker was stuck on the ground). Then they were instructed to put on and calibrate the HMD. Next, participants opened the designated AR app, and underwent the pre-training of a complete book-searching procedure without receiving any hints/feedback. They then used the floating virtual buttons for choosing their specified groups with a voice prompt: "xx groups is selected".
- In the **second phase**, the participants started the first book-searching task by saying "Start one". As the title of a randomly selected book appeared (Figure 1.a) and a built-in timer began (on the top of Figure 1.c), they started searching based on the given title. In this phase, all participants from studies I and II except those from the control groups received either visual/audio hint(s) or the combined set, which pointed out a specified bookcase shelf. The visual hint remained in the real world space regardless of the participants' position. The audio hint was, however, not repeated. Upon successfully finding the book, the participants completed the first search task by saying "Stop task" to stop the timer while the TCT of this task was measured and noted. Next, only the participants from Study II began the review of their gaze playback by articulating "Playback". On finishing, they filled out a NASA Task Load Index (TLX) questionnaire [13], a widely-used assessment tool for perceived workload. The participants were then directed to stand one metre in front of the second test bookcase with saying "Start two" and "Stop task" to start and complete the second task using the same process. The sign of the book found in each task is when the user successfully visually locate the book by seeing its title, and then stop the timer orally. All of the voice commands given by participants to the device were recognized at first try during the formal study because of the quiet environment and they knew how to stop the timer by speech from the pre-training sessions. Finally, they were given a second NASA TLX questionnaire to fill out. All participants were told to unrestrictedly express their feedback after completing the study.

## 5 RESULTS AND ANALYSES

Below we present the results from studies I and II. All participants (n=96) completed the study successfully in finding the correct books at first try in each task with different durations. Thus, no tasking accuracy but the tasking time was harnessed. The dependent variables employed for assessing the tasks were TCT and NASA TLX.

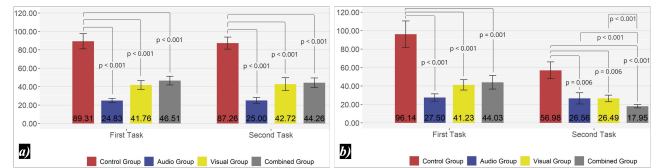


Figure 5: Study I (a: without instant post-task feedback) and Study II (b: with instant post-task feedback): The mean TCT (s) used per task by the four groups. Error bars show mean  $\pm$  standard error (SE).

## 5.1 Study I: Without Instant Post-task Feedback

### 5.1.1 Task Completion Time

We first present the results according to the measured TCT of the two book-searching tasks conducted on task performance. As shown in Figure 5.a, the descriptive statistics show that in the first task the mean TCT for the control group ( $mean = 89.31$ ;  $SD = 29.80$ ) was much longer than for the other three experimental groups: audio group ( $mean = 24.83$ ;  $SD = 8.37$ ), visual group ( $mean = 41.76$ ;  $SD = 17.10$ ), combined group ( $mean = 46.51$ ;  $SD = 17.04$ ). Similarly in the second task, the mean TCT of the control group ( $mean = 87.26$ ;  $SD = 23.61$ ) was still considerably higher than the other groups. Notably, the audio group ( $mean = 25.00$ ;  $SD = 11.70$ ) still had the shortest TCT in comparison with the visual group ( $mean = 42.72$ ;  $SD = 25.36$ ) and the combined group ( $mean = 44.26$ ;  $SD = 18.71$ ). Also notably, the combined group did not have an obvious TCT decrease compared to the other three groups in Study I.

To identify the significant differences of the results, a 2 (task) \* 4 (group) mixed analysis of variance (ANOVA) ( $p = 0.05$ ) was performed given that the normality of the data was affirmed. This analysis was performed using IBM SPSS Statistics, as were the following statistical measurements. Bonferroni-corrected post hoc tests were employed to determine if the pairwise groups were significantly different. We found that all participants did not have statistically significant shorter TCT ( $mean = 49.81$ ;  $SD = 30.51$ ) in the second task compared to the first ( $mean = 50.60$ ;  $SD = 30.61$ ),  $F(1, 44) = 0.183$ ,  $p = 0.671$ . But there were significant main effects of the four groups on the TCT measured ( $F(3, 44) = 21.042$ ,  $p < 0.001$ ). No significant interaction was found between the tasks and the groups ( $F(3, 44) = 0.137$ ,  $p = 0.937$ ). The Bonferroni post hoc tests showed statistical significance between every pairwise comparison of the control group with all the other groups in both tasks. For interpretation, the control group had significantly the most TCTs compared to the audio group ( $p < 0.001$ ), the visual group ( $p < 0.001$ ), and the combined group ( $p < 0.001$ ). No significance was found among the pairs between the audio, visual, and combined groups.

### 5.1.2 NASA TLX

Here, we report NASA TLX answers gathered from the 48 participants from Study I on perceived workload. The descriptive statistics and visualizations of the collective data are shown in Figures 6 and 7, where the mean values with standard errors (SE) are reported. By observing the Total Workload (average value calculated from the six indexes included: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration), we found that the values of this metric from the control group (Total Workload:  $mean = 66.67$ ;  $SD = 17.42$  in the first task;  $mean = 67.5$ ;  $SD = 15.66$  in the second task) were greater than the other three groups for both tasks. Furthermore, participants in the control group yielded the most perceived workload in every index upon every task; in particular, the members of the control group had a substantial lead in Mental Demand ( $mean = 72.67$ ;  $SD = 7.99$  for the first task;  $mean = 70.83$ ;  $SD = 6.72$  for the second task) and Effort ( $mean = 72.08$ ;  $SD = 19.20$  for the first task,  $mean = 74.17$ ;  $SD = 16.81$  for the second task).

Since the NASA TLX values showed variance homogeneity, the 2\*4 mixed ANOVA ( $p = 0.05$ ) was performed to examine the TLX results. The study showed that all participants did not have a significantly different perceived workload ( $mean = 47.73$ ;  $SD = 21.39$ ) in the second task compared to the first ( $mean = 48.14$ ;  $SD = 20.71$ ),  $F(1, 44) = 0.052$ ,  $p = 0.821$ . In contrast, there were significant main effects of the NASA TLX results among the four groups ( $F(3, 44) = 7.591$ ,  $p < 0.001$ ). No significant difference was found on the interaction between the tasks and the groups ( $F(3, 44) = 0.182$ ,  $p = 0.908$ ). The post hoc test showed all three groups were significantly different when compared to the control group in the both tasks. We conducted statistical analysis on the six indexes and Total Workload. However, we only report the results of Total Workload in the following parts of this paper since the significance found on the original six indexes can be referred to in Figures 6 to 9. The participants from the control group felt that there was significantly more perceived workload than those from the audio group ( $mean = 39.09$ ;  $SD = 21.82$ ;  $p = 0.002$ ), the visual group ( $mean = 38.61$ ;  $SD = 25.48$ ;  $p < 0.001$ ), and the combined group ( $mean = 48.19$ ;  $SD = 25.58$ ;  $p = 0.038$ ). Similarly in the second task, the control group ( $mean = 67.50$ ;  $SD = 9.25$ ) had statistical significance between the audio ( $mean = 40.14$ ;  $SD = 19.17$ ;  $p = 0.002$ ), the visual ( $mean = 36.81$ ;  $SD = 19.81$ ;  $p < 0.001$ ), and the combined groups ( $mean = 46.46$ ;  $SD = 22.08$ ;  $p = 0.02$ ). Again, no significant differences between every pairwise comparison of audio, visual, and combined groups were found in both tasks in Study I.

## 5.2 Study II: With Instant Post-task Feedback

### 5.2.1 Task Completion Time

As with Study I, we first considered the results by TCT for the two book-searching tasks performed with the instant post-task feedback in between. As shown in Figure 5.b, the descriptive statistics also show that in the first task the mean TCT for the control group ( $mean = 96.14$ ;  $SD = 49.43$ ) was much longer than for the other three experimental groups: audio group ( $mean = 27.50$ ;  $SD = 13.87$ ), visual group ( $mean = 41.23$ ;  $SD = 19.65$ ), combined group ( $mean = 44.03$ ;  $SD = 25.89$ ). Similarly in the second task, the mean TCT of the control group ( $mean = 56.98$ ;  $SD = 31.59$ ) was still considerably higher than the other groups. However, the combined group ( $mean = 17.95$ ;  $SD = 6.36$ ) had the shortest TCT in comparison with the audio group ( $mean = 22.91$ ;  $SD = 20.89$ ) and the visual group ( $mean = 26.49$ ;  $SD = 12.51$ ). This result differed from that of the first book-searching task.

Likewise, after the normality of the TCT data was determined, the 2\*4 mixed ANOVA ( $p = 0.05$ ) revealed that all participants had significantly shorter TCTs ( $mean = 31.99$ ;  $SD = 25.33$ ) in the second task compared to the first ( $mean = 52.23$ ;  $SD = 40.49$ ),  $F(1, 44) = 13.421$ ,  $p < 0.001$ . The results also showed significant main effects of the four groups on the measured TCT ( $F(3, 44) = 14.390$ ,  $p < 0.001$ ). In addition, there was a significant interaction between the tasks and the groups ( $F(3, 44) = 2.844$ ,  $p = 0.048$ ). The Bonferroni post hoc tests showed significant differences between every pairwise comparison of the control group with all the other groups in the first task. For interpretation, the control group had the longest TCT compared to the audio group ( $p < 0.001$ ), the visual group ( $p < 0.001$ ), and the combined group ( $p = 0.001$ ). No significance was found among the pairs between the audio, visual, and combined groups. Likewise, the second task also showed significant differences between the control group and the other three groups ( $p = 0.006$  for the audio and visual groups,  $p < 0.001$  for the combined group). However, we found that the pairs of the audio-combined ( $p = 0.001$ ) and the visual-combined ( $p = 0.001$ ) had significance as well. The TCT of the combined group was statistically different from the other three groups in the second task.

By observing the time outcomes from both tasks, we found that part of our H1 – individual equipment of visual and audio hints

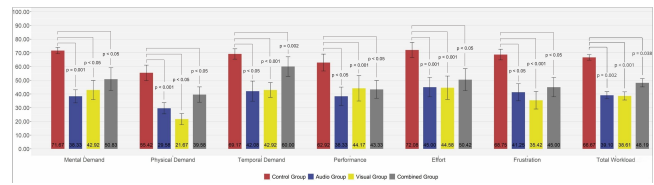


Figure 6: First task in Study I: NASA TLX results of control (left) and three experimental groups. Error bars show mean  $\pm$  standard error.

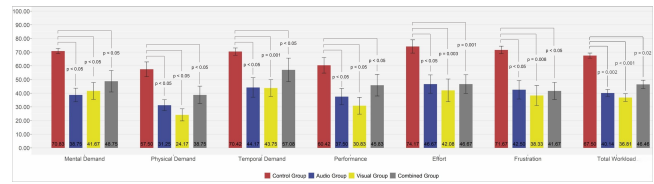


Figure 7: Second task in Study I: NASA TLX results of control (left) and three experimental groups. Error bars show mean  $\pm$  standard error.

have a positive effect in facilitating visual search task performance in AR – conformed to the obtained results. But another part of H1 – the combination of visual and audio hints has greater effect than either does individually in AR searching – was only valid in the second task. Furthermore, the time used in the first task (without instant post-task feedback) exceeded that in the second task (with instant post-task feedback) in every group. Hence, our H2 – Instant post-task feedback contributes to better performance in AR search tasks – was also reflected according to the TCT results.

### 5.2.2 NASA TLX

The NASA TLX answers gathered from study II are then presented. The descriptive statistics and visualizations of the collective data are shown in Figures 8 and 9. By observing the Total Workload, we found that the values of this metric from the control group (Total Workload:  $mean = 65.33$ ,  $SD = 9.46$  in the first task;  $mean = 46.83$ ,  $SD = 7.03$  in the second task) were greater than the other three groups for both tasks. Furthermore, participants in the control group yielded the most perceived workload in every index upon every task; in particular, the members of the control group had an evident lead in Mental Demand ( $mean = 69.17$ ,  $SD = 17.06$ ), Effort ( $mean = 77.08$ ,  $SD = 12.98$ ), and Frustration ( $mean = 67.50$ ,  $SD = 16.39$ ) in the first task. For the second task, the differences were not as remarkable; Mental Demand for the control group ( $mean = 50.00$ ,  $SD = 24.83$ ) retained the high values. We noticed that for the second task almost all of the workload indexes, including total workload in the combined group were lower than the other three groups; the exception was Physical Demand.

The variance homogeneity was once again revealed in the TLX values. Therefore, the result of the 2\*4 mixed ANOVA ( $p = 0.05$ ) showed that all participants had a statistically significantly lesser perceived workload ( $mean = 31.25$ ;  $SD = 11.96$ ) in the second task compared to the first ( $mean = 46.87$ ;  $SD = 13.89$ ),  $F(1, 44) = 45.344$ ,  $p < 0.001$ . There were significant main effects of the NASA TLX results among the four groups ( $F(3, 44) = 6.049$ ,  $p = 0.002$ ), but no significant difference was found on the interaction between the tasks and the groups ( $F(3, 44) = 2.387$ ,  $p = 0.082$ ). For pairwise comparison, the post hoc test showed no significant differences between every pairwise comparison of audio, visual, and combined groups in the first task. However, all the other three groups were significantly different when compared to the control group. That is, the participants from the control group felt that there was significantly more perceived workload than those from the audio group ( $mean =$ ;  $SD = 5.71$   $p = 0.003$ ), the visual group

(mean = 36.50; SD = 10.82;  $p = 0.001$ ), and the combined group (mean = 45.83; SD = 6.97;  $p = 0.021$ ). In the second task, the participants from the control group still possessed significantly more perceived workload compared to members of the audio group (mean = 32.00; SD = 6.69;  $p = 0.003$ ), the visual group (mean = 24.33; SD = 7.31;  $p = 0.02$ ), and the combined group (mean = 21.83; SD = 7.36;  $p = 0.009$ ). Nonetheless, the pairs of the audio-combined ( $p = 0.027$ ) and the visual-combined ( $p = 0.006$ ) also revealed significant differences. That is, The results from the combined group are statistically different from the other groups in the second task. The result means the synthesis of the two modalities of the hints we tested worked better than when there was only one modal hint in the second task, where the gaze playback was provided as the instant post-task feedback in the same context.

It is noted that the non-control groups had an obvious workload decline compared to the control group in each task whilst the combined group had the lowest perceived workload in the second task (H1) in Study II. All the indexes including Total Workload in every group showed considerable decrease for the second task (H2).

### 5.3 Summary of Qualitative Analysis

#### 5.3.1 Hints

Most participants expressed compliments for both the visual and audio hints in the process of searching the books with the AR glasses mounted. The participants mostly stated that they obtained much help from the visual hint in guiding them during the searching procedure. One participant in the visual group from Study I said: *"The arrow was brightly, clearly, and precisely pointing out the correct shelf which guided me finding the book."* Another participant in the visual group from Study II explained: *"It was great to see the visual arrow appear immediately because I was expecting something slow. Also, the arrow sign was big so that I could find the book in a time-saving way."* One participant from the combined group in Study II described the high comprehensibility of the hint: *"It was easy to understand the meaning of the arrow that facilitated the pace of my book searching."* The audio hint was praised mainly due to its unambiguity and clearness. One participant from the audio group in Study I reported: *"the auditory message was an excellent directive in leading me to the target shelf. It was a very clear information with an appropriate rate that I could catch the main content smoothly."* Another participant from the combined group in Study II said: *"The hint was clear and concise which was easy to follow. It took little time to comprehend so that it was helpful for me to set the goal during searching the book."*

There were also some critiques for the two types of hints. One participant from the visual group in Study I said: *"It was a bit difficult to notice there was a visual arrow since it was out of my sight, but it turned out to be working after I realized that. The arrow could be put on the position closer to the centre of the sight."* For the audio hint, one participant from the combined group in Study II suggested *"I would have missed the audio instruction since it was a short message, which needed me to concentrate more on the task. It would be better if the hint was followed by one reminder on the screen, such as "please be aware of the incoming audio hint"."*

#### 5.3.2 Task Feedback

Only the 48 participants in Study II experienced the instant post-task feedback. Most of them gave a positive appraisal of the gaze playback. For example, one participant commented: *"The gaze playback was a good feedback since it displayed to me how I was searching for the book in the first task, which made me get better self-organized on how I tried to locate the book in the second task."* Another participant told us: *"the gaze playback was amazing since it told me do not look at other irrelevant places when searching the second book."* However, one participant also criticized: *"the dot fluctuated so quickly that I could not follow and understand what it*

*intended to tell me. It could be better if the gaze trajectory line was also visualized."*

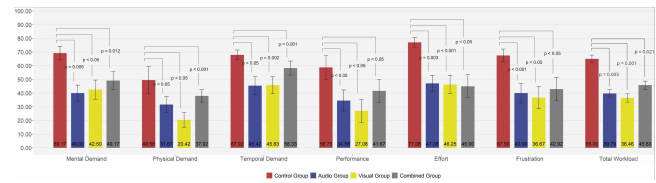


Figure 8: First task in Study II: NASA TLX results of control (left) and three experimental groups. Error bars show mean  $\pm$  standard error.

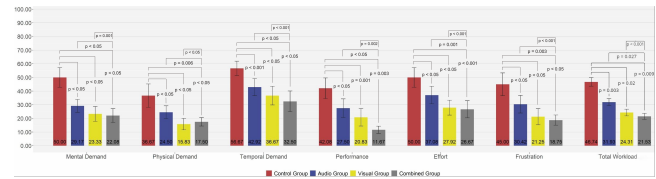


Figure 9: Second task in Study II: NASA TLX results of control (left) and three experimental groups. Error bars show mean  $\pm$  standard error.

Overall, both the visual/audio hints and instant post-task feedback were regarded helpful by all participants. When checking the feedback given by them, we realized that the visual arrow and audio message were playing a dominant role in guiding the participants to the targets, which substantially shortened the task periods and decreased the workload. We also believe that after reviewing the gaze playback, the participants gained more awareness of how to perform and improve the book-searching task.

### 5.4 Findings

By analyzing all the results from the two studies, we found that the TCT and task perceived workload reported gave identical results thus confirming our hypotheses. The visual/audio/combined groups all had a significant difference from the control group in both TCT and NASA TLX, while the two tasks were significantly different when coupling with the instant post-task feedback. In addition, the combined group had significant differences with both visual and audio groups in Study II where the feedback was supplied. We believe the results obtained are representative and can be harnessed to verify our hypotheses. Thus, we rephrase our findings as follows since H1 was partially verified while H2 and was completely affirmed.

- Even though visual and audio hints both have a positive effect in facilitating task performance and reducing perceived workload in AR visual search tasks, the combination of these two hints has a greater effect than either does individually **under the condition that there is instant post-task feedback.**
- Instant post-task feedback through gaze assistance has the capacity to stimulate better performance and reduce perceived workload in the same context.

## 6 DISCUSSION AND LIMITATIONS

### 6.1 Insights

What did we do and achieve in our experiments? We started from two core hypotheses, H1 and H2, with the purpose of discovering the effects of visual/audio hints in either single or combined setups for AR search tasks. We also studied the influence of instant post-task feedback in the same context. To resolve the assumptions, we developed an AR approach engaging two modules of the hints and gaze-assisted instant post-task feedback separately. We

designed a case study for a visual book-searching task which involved the use of a mobile AR headset. An AR app was installed in Microsoft HoloLens 2. We first ran a pre-testing session to assure the functionality of our AR solution, as well as to gain early-stage feedback. After some adjustments, we designed and implemented a more comprehensive, larger-scale user study, where 96 participants were allocated in Study I (n=48) and II (n=48) and placed randomly four groups (control (n=12), audio (n=12), visual (n=12), and combined (n=12)) in each sub-study. The user study was based on the between-subject principle with within-subject factor and the collected results were analyzed correspondingly. Participants were extensively engaged in searching tasks within two distinct bookcases, each containing a diverse range of totally different books, ensuring the adequacy of the study for deriving meaningful results. Moreover, we have taken into account the exposure time of participants to experimental conditions, assuring that it does not compromise the validity of the results. Throughout the entire study and data analysis, **H1** was partially verified, while **H2** was completely verified.

Our proposed solution proves to be advantageous for increasing AR visual search task performance and simultaneously reducing the perceived workload on users. We found that the combination of visual and audio hints can ameliorate both task performance and perceived workload when coupled with instant post-task feedback. We think that participants were mentally self-reflected and guided by the feedback regarding their searching modes during the tasks, as mentioned in qualitative results. A few reported that they improved the task performance by slightly changing the searching mode/route in the second task. Also, the "double guidance" brought by the combination of visual and audio hints positively affected participants after their self-reflection from the post-task feedback. We believe our results are promising, both encouraging further investigations in the field of using directive hints within mobile HMD setups, and incorporating an immediate opportunity for user self reflection and evaluation. The noticeable decrease of TCT and the cognitive task workload was a confirmation of our hypotheses. Beyond the fundamental results obtained, the research implication of our study has high generalizability and convertibility. We aimed to benefit the general AR searching specialization by selecting the visual book-searching task since all the essential elements needed for visual searching were contained in this use case.

## 6.2 Limitations

We acknowledge that our study has some limitations. First, we designed two visual book-searching tasks to be performed using two bookcases that contained different books, eliminating memory as a factor in the second task. Second, The books themselves can serve as the variable elements. The thickness and order of the books might alter the participants' search speeds and workload. The books we chose were generally thick ones, and the participants' searches might have been sped up to some extent because the title displayed on the spine would be larger. Also, the books' positioning may affect participants' searching times since individuals can have varied viewing habits. A book put in the middle, for example, might be more easily noticed by people who are used to searching from the middle; a person who was used to searching from left to right might find a book on the left faster in the first task, while taking longer to find a book located on the right in the second task. Additionally, some books having similar titles might disturb and confuse participants which may lead to a longer or incorrect searching process. Nonetheless, there could potentially exist ambient visual or auditory disruptions (such as the objects except books on the bookcases and unforeseen background noises) that might have a slight impact on the participants.

Third, the inclusion of participants of our study might not be sufficient. A more diverse group of samples could be considered in the future to strengthen the reliability of the research conclusions.

For example, we did not yet consider incorporating non-binary or LGBTQ+ participants [53]. Fourth, we employed solely the NASA TLX as the subjective evaluation measure. Other subjective questionnaires could be utilized to incorporate more diverse subjective aspects, such as satisfaction and pleasure. Fifth, due to the performance sequence, the gaze review only influenced the second task in our study. However, in order to better understand the influence of gaze review on search tasks, there could be another design solution with a counterbalancing factor [5] where the gaze playback could be recorded from the second task then used for the first task.

Finally, there is room for improvement in the AR app, especially for the design of the visual/audio hints. The arrow used in the current app was calibrated to the position of the HMD. Thus, the relative heights of arrows and the bookcases differed due to the different heights of the participants. As a result, an arrow pointing to a shelf might go slightly up or down. Participants in the visual group might perceive higher workload and spend more time aligning the arrow with a specific shelf due to inaccurate calibration. To address this, future work should calibrate the arrow positions with the bookcases rather than the HMD itself. In addition, the visual arrow used in our study led to advisable results but it might be too monotonous. Different design principles of the visual hints should be considered, e.g., not only the arrows but also other types of graphic representations which can visually guide participants. Furthermore, more options of the color rendering, the size, and the positioning of the visual hints should be discussed and compared to make the proof-of-concept more robust.

## 7 CONCLUSIONS AND FUTURE WORK

In this paper, we have described an AR approach for visual search tasks with supported visual/audio hints and gaze-assisted instant post-task feedback based on mobile HMD. Previous research had given us an incentive to investigate the effect of hints and feedback. Based on our hypotheses, we designed and conducted a case study of visual book-searching where the gaze playback acted as the instant post-task feedback based. The experimental procedure consisted of a comprehensive user study (n=96) with two comparative sub-studies. The resulting analysis was founded mainly on collected NASA TLX answers with TCT measurement as a preliminary analytic metric. We partially verified **H1**: both visual and audio hints have a positive effect in facilitating task performance. The combination of the two hints works better than either individually, under the condition that there is instant post-task feedback. **H2** was completely verified: instant post-task feedback has the capacity to bring about better performance. We pointed out the high generalizability and convertibility of our research output in making advantage of the general AR searching processes.

Our research is novel because it fills the gaps of combining visual and audio hints in the same AR context, the integration of feedback and gaze assistance. More importantly, it not only focuses on the AR hints and instant post-task feedback, but places them in the same context and generates meaningful conclusions for universal AR visual search tasks. In future work, we will make our app more adjustable for human eyes and thereby user-friendly by shortening users' adaptation time. We will also improve our system by adding adaptive support which can react to users' gaze by giving additional help to users focusing too long on the wrong parts of the environment. Next, we intend to involve more types of visual/audio hints, as well as more modalities of feedback. Meanwhile, more demographic information will be collected and more effort will be made to involve LGBTQ+ participants. We foresee carrying out a more inclusive user study with more metrics for evaluation. We hope that our research will have an impact on industry-based AR setups and configurations empowering people's capabilities and levels of efficiency in general AR search tasks.



## ACKNOWLEDGMENTS

This work was partly supported by the Norges Forskningsrad (309339, 314578), MediaFutures user partners and Universitetet i Bergen.

## REFERENCES

- [1] A. Ajanki, M. Billinghurst, H. Gamper, T. Järvenpää, M. Kandemir, S. Kaski, M. Koskela, M. Kurimo, J. Laaksonen, K. Puolamäki, et al. An augmented reality interface to contextual information. *Virtual reality*, 15(2-3):161–173, 2011. 3
- [2] F. Anderson, T. Grossman, J. Matejka, and G. Fitzmaurice. Youmove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pp. 311–320, 2013. 2
- [3] S. Arevalo Arboleda, F. Rücker, T. Dierks, and J. 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*, pp. 1–14, 2021. 2, 3
- [4] R. Behringer, J. Park, and V. Sundareswaran. Model-based visual tracking for outdoor augmented reality applications. In *Proceedings. international symposium on mixed and augmented reality*, pp. 277–322. IEEE, 2002. 2
- [5] J. V. Bradley. Complete counterbalancing of immediate sequential effects in a latin square design. *Journal of the American Statistical Association*, 53(282):525–528, 1958. 8
- [6] Y. Cao, T. Wang, X. Qian, P. S. Rao, M. Wadhawan, K. Huo, and K. Ramani. Ghostar: A time-space editor for embodied authoring of human-robot collaborative task with augmented reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 521–534, 2019. 4
- [7] G. Charness, U. Gneezy, and M. A. Kuhn. Experimental methods: Between-subject and within-subject design. *Journal of economic behavior & organization*, 81(1):1–8, 2012. 2
- [8] M. Cidota, S. Lukosch, D. Datcu, and H. Lukosch. Comparing the effect of audio and visual notifications on workspace awareness using head-mounted displays for remote collaboration in augmented reality. *Augmented Human Research*, 1(1):1–15, 2016. 2, 3
- [9] F. Clemente, S. Dosen, L. Lonini, M. Markovic, D. Farina, and C. Cipriani. Humans can integrate augmented reality feedback in their sensorimotor control of a robotic hand. *IEEE Transactions on Human-Machine Systems*, 47(4):583–589, 2016. 2, 3
- [10] P. Contreras, D. Chimbo, A. Tello, and M. Espinoza. Semantic web and augmented reality for searching people, events and points of interest within a university campus. In *2017 XLIII Latin American Computer Conference (CLEI)*, pp. 1–10. IEEE, 2017. 1, 2
- [11] C. Curatu, H. Hua, and J. Rolland. Projection-based head-mounted display with eye-tracking capabilities. In *Novel Optical Systems Design and Optimization VIII*, vol. 5875, p. 58750J. International Society for Optics and Photonics, 2005. 3
- [12] C. Gebhardt, B. Hecox, B. van Opheusden, D. Wigdor, J. Hillis, O. Hilliges, and H. Benko. Learning cooperative personalized policies from gaze data. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 197–208, 2019. 2
- [13] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988. 5
- [14] Y.-Y. Huang and M. Menozzi. Effects of viewing distance and age on the performance and symptoms in a visual search task in augmented reality. *Applied Ergonomics*, 102:103746, 2022. 1
- [15] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billinghurst. Pinpointing: Precise head-and eye-based target selection for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2018. 2
- [16] M. Lankes and B. Stiglbauer. Gazeair: Mobile gaze-based interaction in the context of augmented reality games. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 397–406. Springer, 2016. 2
- [17] J.-Y. Lee, H.-M. Park, S.-H. Lee, T.-E. Kim, and J.-S. Choi. Design and implementation of an augmented reality system using gaze interaction. In *2011 International Conference on Information Science and Applications*, pp. 1–8. IEEE, 2011. 1
- [18] R. W. Lindeman, H. Noma, and P. G. De Barros. Hear-through and mic-through augmented reality: Using bone conduction to display spatialized audio. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 173–176. IEEE, 2007. 2
- [19] C. Liu, S. Huot, J. Diehl, W. Mackay, and M. Beaudouin-Lafon. Evaluating the benefits of real-time feedback in mobile augmented reality with hand-held devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2973–2976, 2012. 3
- [20] J. Looser, M. Billinghurst, R. Grasset, and A. Cockburn. An evaluation of virtual lenses for object selection in augmented reality. In *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, pp. 203–210, 2007. 3
- [21] F. Lu, S. Davari, L. Lisle, Y. Li, and D. A. Bowman. Glanceable ar: Evaluating information access methods for head-worn augmented reality. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 930–939. IEEE, 2020. 3
- [22] K. Lyons, M. Gandy, and T. Starner. Guided by voices: An audio augmented reality system. Georgia Institute of Technology, 2000. 1, 2
- [23] A. Marquardt, C. Trepkowski, T. D. Eibich, J. Maiero, and E. Kruijff. Non-visual cues for view management in narrow field of view augmented reality displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 190–201. IEEE, 2019. 2, 4
- [24] A. Mulloni, H. Seichter, and D. Schmalstieg. User experiences with augmented reality aided navigation on phones. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, pp. 229–230. IEEE, 2011. 3
- [25] K. Murakami, R. Kiyama, T. Narumi, T. Tanikawa, and M. Hirose. Poster: A wearable augmented reality system with haptic feedback and its performance in virtual assembly tasks. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 161–162. IEEE, 2013. 3
- [26] E. Nonino, J. Gisler, V. Holzwarth, C. Hirt, and A. Kunz. Subtle attention guidance for real walking in virtual environments. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 310–315. IEEE, 2021. 2
- [27] A. Nowak, Y. Zhang, A. Romanowski, and M. Fjeld. Augmented reality with industrial process tomography: To support complex data analysis in 3d space. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, pp. 56–58, 2021. 1
- [28] S. Oney, N. Rodrigues, M. Becher, T. Ertl, G. Reina, M. Sedlmair, and D. Weiskopf. Evaluation of gaze depth estimation from eye tracking in augmented reality. In *ACM Symposium on Eye Tracking Research and Applications*, pp. 1–5, 2020. 2
- [29] V. Paelke. Augmented reality in the smart factory: Supporting workers in an industry 4.0. environment. In *Proceedings of the 2014 IEEE emerging technology and factory automation (ETFA)*, pp. 1–4. IEEE, 2014. 4
- [30] H. M. Park, S. H. Lee, and J. S. Choi. Wearable augmented reality system using gaze interaction. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 175–176. IEEE, 2008. 2, 3
- [31] N. Pathmanathan, M. Becher, N. Rodrigues, G. Reina, T. Ertl, D. Weiskopf, and M. Sedlmair. Eye vs. head: Comparing gaze methods for interaction in augmented reality. In *ACM Symposium on Eye Tracking Research and Applications*, pp. 1–5, 2020. 3
- [32] N. Petersen, A. Pagani, and D. Stricker. Real-time modeling and tracking manual workflows from first-person vision. In *2013 IEEE International symposium on mixed and augmented reality (ISMAR)*, pp. 117–124. IEEE, 2013. 2
- [33] V. Peysakhovich, O. Lefrançois, F. Dehais, and M. Causse. The neuroergonomics of aircraft cockpits: The four stages of eye-tracking integration to enhance flight safety. *Safety*, 4(1), 2018. doi: 10.3390/safety4010008 3
- [34] K. Pfeuffer, Y. Abdrabou, A. Esteves, R. Rivu, Y. Abdelrahman, S. Meitner, A. Saadi, and F. Alt. Arattention: A design space for gaze-

- adaptive user interfaces in augmented reality. *Computers & Graphics*, 95:1–12, 2021. 1
- [35] R. Piening, K. Pfeuffer, A. Esteves, T. Mittermeier, S. Prange, P. Schröder, and F. Alt. Looking for info: Evaluation of gaze based information retrieval in augmented reality. In *IFIP Conference on Human-Computer Interaction*, pp. 544–565. Springer, 2021. 1, 2
- [36] J. Qian, J. Ma, X. Li, B. Attal, H. Lai, J. Tompkin, J. F. Hughes, and J. Huang. Portal-ble: Intuitive free-hand manipulation in unbounded smartphone-based augmented reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 133–145, 2019. 1
- [37] A. Rafiq and B. Ahsan. Secure and dynamic model for book searching on cloud computing as mobile augmented reality. *International Journal of Modern Education & Computer Science*, 6(1), 2014. 2
- [38] C. Reardon, K. Lee, and J. Fink. Come see this! augmented reality to enable human-robot cooperative search. In *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pp. 1–7. IEEE, 2018. 1
- [39] A. Riegler, B. Aksoy, A. Riener, and C. Holzmann. Gaze-based interaction with windshield displays for automated driving: Impact of dwell time and feedback design on task performance and subjective workload. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 151–160, 2020. 3
- [40] R. Rivu, Y. Abdrabou, K. Pfeuffer, A. Esteves, S. Meitner, and F. Alt. Stare: Gaze-assisted face-to-face communication in augmented reality. In *ACM Symposium on Eye Tracking Research and Applications*, pp. 1–5, 2020. 1, 3
- [41] P. Sasikumar, L. Gao, H. Bai, and M. Billingham. Wearable remotefusion: A mixed reality remote collaboration system with local eye gaze and remote hand gesture sharing. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 393–394. IEEE, 2019. 3
- [42] N. Sawhney and C. Schmandt. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM transactions on Computer-Human interaction (TOCHI)*, 7(3):353–383, 2000. 1, 2
- [43] L. Scolere and L. Malinin. The building that teaches: Exploring augmented reality affordances in academic incubators. *Journal of Interior Design*, 2022. 2
- [44] A. Seeliger, G. Merz, C. Holz, and S. Feuerriegel. Exploring the effect of visual cues on eye gaze during ar-guided picking and assembly tasks. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 159–164. IEEE, 2021. 2
- [45] F. S. Shaikh. Augmented reality search to improve searching using augmented reality. In *2021 6th International Conference for Convergence in Technology (I2CT)*, pp. 1–5. IEEE, 2021. 1
- [46] M. Sousa, J. Vieira, D. Medeiros, A. Arsenio, and J. Jorge. Sleeve: Augmented reality for rehabilitation using realtime feedback. In *Proceedings of the 21st international conference on intelligent user interfaces*, pp. 175–185, 2016. 2
- [47] V. Sundareswaran, K. Wang, S. Chen, R. Behringer, J. McGee, C. Tam, and P. Zahorik. 3d audio augmented reality: implementation and experiments. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, pp. 296–297. IEEE, 2003. 2
- [48] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980. 1
- [49] C. Trepkowski, D. Eibich, J. Maiero, A. Marquardt, E. Kruijff, and S. Feiner. The effect of narrow field of view and information density on visual search performance in augmented reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 575–584. IEEE, 2019. 2
- [50] J. Van Dam, A. Krasne, and J. 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)*, pp. 201–204. IEEE, 2020. 2
- [51] J. Vieira, M. Sousa, A. Arsenio, and J. Jorge. Augmented reality for rehabilitation using multimodal feedback. In *Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques*, pp. 38–41, 2015. 2
- [52] B. Volmer, J. Baumeister, S. Von Itzstein, I. Bornkessel-Schlesewsky, M. Schlesewsky, M. Billingham, and B. H. Thomas. A comparison of predictive spatial augmented reality cues for procedural tasks. *IEEE transactions on visualization and computer graphics*, 24(11):2846–2856, 2018. 1
- [53] A. M. Walker and M. A. DeVito. "more gay" fits in better": Intra-community power dynamics and harms in online lgbtq+ spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2020. 8
- [54] S. White, L. Lister, and S. Feiner. Visual hints for tangible gestures in augmented reality. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 47–50. IEEE, 2007. 1, 2
- [55] D. Wolf, D. Besserer, K. Sejunaite, M. Riepe, and E. Rukzio. care: An augmented reality support system for dementia patients. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*, pp. 42–44, 2018. 3
- [56] J. Wolf, Q. Lohmeyer, C. Holz, and M. Meboldt. Gaze comes in handy: Predicting and preventing erroneous hand actions in ar-supported manual tasks. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 166–175. IEEE, 2021. 2
- [57] J. Yip, S.-H. Wong, K.-L. Yick, K. Chan, and K.-H. Wong. Improving quality of teaching and learning in classes by using augmented reality video. *Computers & Education*, 128:88–101, 2019. 2
- [58] M. Zahriri, C. A. Nelson, D. Oleynikov, and K.-C. Siu. Evaluation of augmented reality feedback in surgical training environment. *Surgical Innovation*, 25(1):81–87, 2018. 2, 3
- [59] Y. Zhang, A. Nowak, G. Rao, A. Romanowski, and M. Fjeld. Is industrial tomography ready for augmented reality? a need-finding study of how augmented reality can be adopted by industrial tomography experts. In *International Conference on Human-Computer Interaction*. Springer, 2023. 2
- [60] Y. Zhang, A. Nowak, A. Romanowski, and M. Fjeld. An initial exploration of visual cues in head-mounted display augmented reality for book searching. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*, pp. 273–275, 2022. 2
- [61] Y. Zhang, A. Nowak, A. Romanowski, and M. Fjeld. On-site or remote working?: An initial solution on how covid-19 pandemic may impact augmented reality users. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces*, pp. 1–3, 2022. 1
- [62] Y. Zhang, A. Nowak, A. Romanowski, and M. Fjeld. Virtuality or physicality? supporting memorization through augmented reality gamification. In *Companion Proceedings of the 2023 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, pp. 53–58, 2023. 1
- [63] Y. Zhang, A. Omrani, R. Yadav, and M. Fjeld. Supporting visualization analysis in industrial process tomography by using augmented reality—a case study of an industrial microwave drying system. *Sensors*, 21(19):6515, 2021. 1
- [64] Y. Zhang, Y. Xuan, R. Yadav, A. Omrani, and M. Fjeld. Playing with data: An augmented reality approach to interact with visualizations of industrial process tomography. *arXiv preprint arXiv:2302.01686*, 2023. 1
- [65] Y. Zhang, R. Yadav, A. Omrani, and M. Fjeld. A novel augmented reality system to support volumetric visualization in industrial process tomography. In *Proceedings of the 2021 Conference on Interfaces and Human Computer Interaction*, pp. 3–9, 2021. 1
- [66] S. Zhao, X. Xiao, Q. Wang, X. Zhang, W. Li, L. Soghier, and J. Hahn. An intelligent augmented reality training framework for neonatal endotracheal intubation. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 672–681. IEEE, 2020. 2
- [67] Y. Zhao, S. Szpiro, J. Knighten, and S. Azenkot. CueSee: exploring visual cues for people with low vision to facilitate a visual search task. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 73–84, 2016. 2
- [68] Z. Zhu, V. Branzoi, M. Wolverson, G. Murray, N. Vitovitch, L. Yarnall, G. Acharya, S. Samarasekera, and R. Kumar. Ar-mentor: Augmented reality based mentoring system. In *2014 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 17–22. IEEE, 2014. 2