Glanceable AR: Towards a Pervasive and Always-On Augmented Reality Future

Feiyu Lu

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Computer Science & Applications

Doug A. Bowman, Chair
Joseph L. Gabbard
Wallace Lages
Sang Won Lee
Yalong Yang
Steven K. Feiner

May 5th, 2023
Blacksburg, Virginia

Keywords: Pervasive Augmented Reality, Input & Interactions, Mobile Computing,
Glanceable Displays
Copyright 2023, Feiyu Lu
Glanceable AR: Towards a Pervasive and Always-On Augmented Reality Future

Feiyu Lu

(ABSTRACT)

Augmented reality head-worn displays (AR HWDs) have the potential to assist personal computing and the acquisition of everyday information. With advancements in hardware and tracking, these devices are becoming increasingly lightweight and powerful. They could eventually have the same form factor as normal pairs of eyeglasses, be worn all-day, overlaying information pervasively on top of the real-world anywhere and anytime to continuously assist people’s tasks. However, unlike traditional mobile devices, AR HWDs are worn on the head and always visible. If designed without care, the displayed virtual information could also be distracting, overwhelming, and take away the user’s attention from important real-world tasks. In this dissertation, we research methods for appropriate information displays and interactions with future all-day AR HWDs by seeking answers to four questions: (1) how to mitigate distractions of AR content to the users; (2) how to prevent AR content from occluding the real-world environment; (3) how to support scalable on-the-go access to AR content; and (4) how everyday users perceive using AR systems for daily information acquisition tasks. Our work builds upon a theory we developed called Glanceable AR, in which digital information is displayed outside the central field of view of the AR display to minimize distractions, but can be accessed through a quick glance. Through five projects covering seven studies, this work provides theoretical and empirical knowledge to prepare us for a pervasive yet unobtrusive everyday AR future, in which the overlaid AR information is easily accessible, non-invasive, responsive, and supportive.
Glanceable AR: Towards a Pervasive and Always-On Augmented Reality Future

Feiyu Lu

(GENERAL AUDIENCE ABSTRACT)

Augmented reality (AR) refers to a technology in which digital information is overlaid on the real-world environment. This provides great potential for everyday uses, because users can view and interact with digital apps anywhere and anytime even when physical screens are unavailable. However, depending on how the digital information is displayed, it could quickly occupy the user’s view, block the real-world environment, and distract or overwhelm users. In this dissertation work, we research ways to deliver and interact with virtual information displayed in AR head-worn displays (HWDs). Our solution centers around the Glanceable AR concept, in which digital information is displayed in the periphery of users’ views to remain unobtrusive, but can be accessed through a glance when needed. Through empirical evaluations, we researched the feasibility of such solutions, and distilled lessons learned for future deployment of AR systems in people’s everyday lives.
Dedication

To everyone I met along the way.
Acknowledgments

It has been a fulfilling five-year Ph.D. journey. I would not have achieved any of these without the continuous and overwhelming support from my advisor, mentors, family, and friends.

First and foremost, I would like to thank my advisor, Prof. Doug A. Bowman. Working with you for the past few years is one of the best and most life-changing experience that ever occurred to me. Not only that I learned how to establish myself as a competent academic researcher, more importantly, I learned how to work with others with a generous, warm, understanding, and humble heart. Starting from Aug 7th 2018, every single one of our discussions has been a enjoyable learning and growing experience. I could not thank you more for nurturing me, for being an all-around role model, for your unwavering support and benevolent care throughout my research journey. It has always been my greatest pleasure, honour, and privilege to have an advisor like you. I will for sure miss our weekly meetings which gave birth to all the sparkling research ideas. My goal is that one day, I will have mentees and students who think about me the same way as how I think about you.

I would like to thank my mentors during my internships at Reality Labs Research, Yan Xu, Missie Smith, Sophie Kim, Laird Malamed, Reid Pinkham, and Michael Shvartsman. Thank you for the countless meetings and discussions, for sharing with me your expertise and resources, and for guiding me on how to become a successful industrial researcher. I gained so much during my amazing time there. I would also like to thank my undergraduate advisor, Prof. Hai-Ning Liang. Thank you for introducing me to the magnificent VR/AR research world. Without you, I would have never realized where my true passions are.
I would like to thank my twin sister Siyu Lu and my parents Mengxue Xu, Ye Lu. Thank you for your love and support. It were your encouragements that made me through the hard times. I would not have come this far if it was not that you were always there for me when I needed someone to talk to.

I would like to thank my Ph.D. committee members, Prof. Joseph L. Gabbard, Prof. Wallace Lages, Prof. Sang Won Lee, Prof. Yalong Yang, and Prof. Steven K. Feiner. Thank you for your time, guidance, encouragement, support, and for raising the intriguing and thought-provoking questions that made my work tons better.

I would like to thank all my labmates in the 3D Interaction Group, Shakiba Davari, Leonardo Pavanatto, Yuan Li, Lei Zhang, Lee Lisle, Ibrahim Tahmid, Jerald Thomas, Alexander Giovannelli, Logan Lane, Francielly Rodrigues, Cory Ilo, Kylie Davidson, and Cherelle Connor. Thank you for piloting my studies and being great research companions. It was the presence of all of you that made my journey fulfilled, accompanied, motivated, and fun.

I would like to thank all the administrative staff in the Department of Computer Science, the Institute for Creativity, Arts, and Technology, and the Center for Human-Computer Interaction at Virginia Tech, Sharon Kinder-Potter, Cliff Shaffer, Andrea Kavanaugh, Holly Williams, Teresa Hall, Melissa Wyers, George Hardebeck, and Cal Ribbens. Thank you for the administrative and technical support.

Last but not least, I would like to thank the Office of Naval Research for their generous support on my research.
Contents

List of Figures xvi

List of Tables xxv

1 Introduction 1

1.1 Motivations 1

1.2 Research Questions 3

1.3 Methods and Scope 5

1.3.1 The Glanceable AR Paradigm 5

1.3.2 RQ1: Mitigating Distractions 6

1.3.3 RQ2: Resolving Occlusions 7

1.3.4 RQ3: Supporting Scalable Transitions 7

1.3.5 RQ4: Real-World Evaluations 9

1.4 Contributions 10

2 Review of Literature 11

2.1 The Evolving Eras of Computing 11

2.2 Pervasive and Everyday Augmented Reality 13

2.3 General-Purpose Information Displays in AR 14
3 Designing for Unobtrusive and Non-Distracting Information Access

3.1 Overview

3.2 Glanceable AR: Initial Design and Concept

   3.2.1 Glanceable AR

   3.2.2 Initial Design Principles of Glanceable AR

   3.2.3 Glanceable AR Interfaces

3.3 Evaluations of Glanceable AR in Walking Scenarios

   3.3.1 Research Questions

   3.3.2 Hypothesis

   3.3.3 Participants

   3.3.4 Tasks

   3.3.5 Measures

   3.3.6 Procedures

   3.3.7 Results

   3.3.8 Summary of Findings

   3.3.9 Discussions
3.4 Conclusions ................................................................. 49

4 Resolving Occlusion Issues of AR Content ............................... 51

4.1 Overview ........................................................................ 51

4.2 Validating Occlusion Issues of AR Content ......................... 53

4.2.1 Interface Conditions .................................................. 53

4.2.2 Tasks ........................................................................ 53

4.2.3 Participants & Procedures ............................................ 54

4.2.4 Results & Discussion .................................................. 55

4.3 Techniques for Rapid Activation of Glanceable Content .......... 57

4.3.1 Design Considerations ................................................ 57

4.3.2 Activation Techniques ................................................. 59

4.4 Evaluations in Stationary and Mobile Contexts .................... 63

4.4.1 Research Goals ......................................................... 63

4.4.2 Participants .............................................................. 64

4.4.3 Tasks ...................................................................... 64

4.4.4 Apparatus ............................................................... 67

4.4.5 Measures ................................................................. 69

4.4.6 Hypotheses ............................................................. 69

4.4.7 Procedure ............................................................... 70
4.4.8 Results .............................................................. 71
4.4.9 Summary of Findings ........................................... 77
4.4.10 Discussion ........................................................... 78
4.4.11 Design Recommendations ................................... 80
4.5 Conclusions ............................................................. 81

5 Supporting Spatial Transitions with AR Interfaces  82

5.1 Overview ............................................................... 82
5.2 Modelling User Interactions During Spatial Transition  86
  5.2.1 User performance modeling and scalability in Interfaces  86
  5.2.2 Research Questions ............................................. 87
  5.2.3 Interface Conditions ............................................. 87
  5.2.4 Hypothesis ......................................................... 89
  5.2.5 Tasks ................................................................. 92
  5.2.6 Participants & Procedures ..................................... 96
  5.2.7 Results ............................................................. 97
  5.2.8 Discussion ........................................................ 101
  5.2.9 Conclusions ....................................................... 106
5.3 Evaluating Realistic Transitions In-Situ .................... 107
  5.3.1 Research Questions ............................................. 107
5.3.2 Interfaces ......................................................... 108
5.3.3 System Implementation ........................................ 110
5.3.4 A Contextual Framework and Hypothesis .................... 110
5.3.5 Scenarios & Tasks .............................................. 115
5.3.6 Results .......................................................... 121
5.3.7 Discussion & Design Implications ............................. 130
5.3.8 Design Recommendations ...................................... 133
5.4 Limitations & Future Work ........................................ 135
5.5 Conclusions ....................................................... 136

6 Real-world Deployments of Glanceable AR ..................... 137

6.1 Overview ........................................................ 137

6.2 Surveying Public Opinions About a Glanceable AR Video Prototype .................. 139
   6.2.1 The video prototype .......................................... 139
   6.2.2 Design rationale .............................................. 141
   6.2.3 Research Goals .............................................. 143
   6.2.4 Experiment .................................................. 144
   6.2.5 Participants ............................................... 145
   6.2.6 Results ..................................................... 146
   6.2.7 Discussion .................................................. 150
6.3 Capstone Implementation I: Initial Perceptions in Authentic Everyday Uses.

6.3.1 Working Prototype .................................................... 151
6.3.2 Improvements ............................................................ 153
6.3.3 Research Goals ........................................................... 156
6.3.4 Participants ............................................................... 156
6.3.5 Experiment Procedures ............................................... 157
6.3.6 Results ................................................................. 157
6.3.7 Summary of Findings ................................................ 162
6.3.8 Discussion ............................................................. 163
6.3.9 Conclusion ............................................................. 165

6.4 Capstone Implementation II: Evaluations of a feature-rich system ....... 166

6.4.1 Capstone II Design & Implementations ................................. 166
6.4.2 Research Questions ................................................... 170
6.4.3 Participants & Procedure ............................................ 170
6.4.4 Measures ............................................................... 171
6.4.5 Participants ............................................................... 171
6.4.6 Results ................................................................. 172
6.4.7 Summary of Findings ................................................ 176
6.4.8 Discussion ............................................................. 177
6.4.9 Limitations and Future Work ....................................... 179
A.3 Semi-structured Interview Scriptss ............................................. 236
A.4 IRB Approval Letter ................................................................. 236
A.5 Video Demonstrations .............................................................. 239

Appendix B Resolving Occlusion Issues of AR Content 240

B.1 Preliminary Study ................................................................. 240
  B.1.1 Informed Consent Form ................................................... 240
  B.1.2 Questionnaires ............................................................... 245
B.2 Primary Study .......................................................... 261
  B.2.1 Informed Consent Form ................................................... 261
  B.2.2 Questionnaires ............................................................... 266
  B.2.3 Semi-structured Interview Scriptss ..................................... 269
B.3 IRB Approval Letter ................................................................. 296
B.4 Video Demonstrations .............................................................. 299

Appendix C Supporting Spatial Transitions with AR Interfaces 300

C.1 Modelling User Interactions During Spatial Transition .......... 300
  C.1.1 Informed Consent Form ................................................... 300
  C.1.2 Questionnaires & Interviews ............................................. 305
C.2 Evaluating Realistic Transitions In-Situ ................................. 315
  C.2.1 Informed Consent Form ................................................... 315
C.2.2 Questionnaires ........................................................... 320

C.2.3 Semi-structured Interview Scriptss ................................. 339

C.3 IRB Approval Letter ...................................................... 341

C.4 Video Demonstrations ................................................... 344

Appendix D Real-world Deployments of Glanceable AR 345

D.1 Surveying Public Opinions About a Glanceable AR Video Prototype .... 345
   D.1.1 Informed Consent Form & Questionnaire ...................... 345
   D.1.2 Interviews .............................................................. 378
   D.1.3 IRB Approval Letter ............................................... 380
   D.1.4 Video Demonstrations ............................................. 383

D.2 Capstone Implementation I & II .................................... 383
   D.2.1 Informed Consent Form ......................................... 383
   D.2.2 Questionnaires ...................................................... 388
   D.2.3 IRB Approval Letter ............................................. 406
   D.2.4 Video Demonstrations ............................................. 409
List of Figures

1.1 An illustration of the Glanceable AR concept: (a) virtual content resides at the periphery in low LoD to stay unobtrusive, but (b) can be easily accessed through a glance by the users when needed (Credit: Lei Zhang - leiz@vt.edu). 7

1.2 An illustration of the structure and organization of this dissertation work. 8

2.1 The three eras of modern computing (Roy Want, 2010 [156]). 12

2.2 (a) An example scenario of pervasive AR, in which a virtual calendar app pops in front of the user when needed for quick access and modifications to schedules (Video: Adobe Aero Max - https://youtu.be/IutVq6cPBX0); (b) A envisioned provocative future of AR, in which AR content takes over the real-world and consumes the cognitive bandwidth of the users (Video: Hyperreality by Keiichi Matsuda - https://youtu.be/YJg02ivYzSs). 14

3.1 Three Glanceable AR interfaces proposed in this section: (a): Eye-Glance (EG) interface (HUD), in which virtual content resides at the edge of FoV and is fixed to users’ head; (b) Head-Glance (HG) 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 (GS interface, in which virtual content can be summoned into FoV by gazing at the periphery for 0.5 second.). 27
3.2 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.

3.3 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.

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

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

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

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

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

4.1 The four interfaces in the first study with information at the periphery: (a) HG-Full; (b) HUD-Full; (c) HG-Icon; and (d) HUD-Icon.
4.2 7-point scale ratings on four statements (a-d). 56

4.3 (a) When the user is looking in the direction of minimized virtual content, we propose five interaction techniques to activate the virtual information: (b) Fixation-Glance, in which users converge their gaze at the depth of the content; (c) Head-Depth, in which users lean backward three centimeters; (d) Hand-Overlay, in which users put their hand slightly behind the virtual content; (e) Blink, in which users blink the eye twice within one second; and (f) Dwell, in which users maintain their gaze on the virtual content for one second. 59

4.4 Illustrations of the two task contexts: (a) the sitting context; (b) the walking context. 68

4.5 (a-b) The time it took for participants to answer the questions in the virtual content/real-world for the (a) sitting task, (b) walking task; (c-d) the SUS score for the five interface conditions in the (c) sitting task and (d) walking task; and (e-f) Participants’ response in 7-point Likert scale to the question “How do you like using the interface to access information in the virtual content/real-world” for the (e) sitting task and (f) walking task (±S.E.). 73

4.6 (a-b) The ranking of each interface under (a) sitting and (b) walking task; (c-d) The false activation rates of virtual content for (c) sitting task (d) walking task (±S.E.). 74

4.7 Social acceptance rate in percentage for (a) locations and (b) audiences. 76

5.1 Example of app suggestions on current mobile phone interfaces. 84
5.2 An illustration of a transition-aware interface, in which the system automatically brings a subset of all AR applications that are relevant to users' task space after the user transitions to a new space to seamlessly support their task activities. Left: in the kitchen environment, the Recipe, Fitness, Notes, Clock, Weather, and Video apps were opened; Right: when the user moves to the office, the system brings the Note and Weather app, and automatically opens the Email, Calendar, Message apps to be ready to assist with productivity tasks, as well as the Reading List and Plant apps that are relevant to the office environment. Illustration by Lei Zhang - leiz@vt.edu

5.3 Example scenarios of using the three conditions: (a) the user is at the living room with a bunch of AR apps opened and placed around. He wants to head to the kitchen for some tasks; (b) None: the user arrives at the kitchen. Since it is a different environment, users loses complete access to the apps in the living room, thus having to use a menu to re-open the AR apps that he needs; (c) Some: the user arrives at the kitchen. The system opens the kitchen-related applications automatically for the users (e.g, recipe, timer, shopping list, calories). However, the user would still need to use the menu if he wants to open other apps; (c) All: the system automatically opens all AR applications and places them around in the environment, no matter needed or not.

5.4 The hypothesized correlation between time it takes for the user to find the target app and the total number of app in each interface condition.
5.5 The task scenario, in which two “locations” were simulated by two squares A and B in AR. Participants were asked to move from one square to another and find a target AR application, simulating a typical transition scenario in real-world scenarios.

5.6 (a) The system highlights four apps for the users to click through before transition simulating prior knowledge with the apps; (b) the system would then ask the user to find one of the four apps after transition, an icon appears on the controller indicating the target app to be find at the new location; (c) in the None condition, the users always need to use a menu to reopen the app they need; (d) the user clicks on the opened app using the controller to finish the trial; (e) in the Some condition, the system always brings the first six apps with the users to the new location; (f) in the All condition, the system brings all apps with the users to the new location. In all the conditions, the content and layout of the apps remain unchanged before and after transition in a single trial, but is randomized between trials.

5.7 The time required for participants to retrieve a target app under each interface condition with different number of apps.

5.8 (a) Average time (Y-axis) it took for each condition to find the target app when different numbers of apps are present (X-axis). Strong linear relationships were indicated for None Some-Menu; and All conditions (bars indicate 95% confidence intervals). (b) The ranking of the three interface conditions by the number of participants (Y-axis).

5.9 The average ratings for easiness and workload of the task under each interface condition.
5.10 The overall Some performance when the probability of the needed app being one of the six suggested apps are randomly chosen.

5.11 (a-b) The user raises the hand to summon the menu, the currently opened/closed apps were shown as opaque/transparent icons, and the system automated apps have a blue asterisk mark on the side; (c) the user gazes at the icon for 0.5 second to open an app manually a white boarder graduate appears around the icon as visual feedback for the dwell interaction (the icons in the menu are alphabetically ordered for easy search); (d) the icon turns opaque, indicating that the app has been opened in space; (e) For Some, the icons of the opened apps were shown on the top of the field of view.

5.12 The design of the menu to make users clearly aware of (1) which app is opened/closed in the current location; (2) which app is opened by the system automatically for the current location (only in the Some condition indicated by a blue star); and (3) which app is pinned by the user and will be automatically opened in the next locations (indicated by a pin icon). The apps are arranged alphabetically to reduce the visual search effort.

5.13 Deployments of the AR apps (i.e., the All interface) in the three physical room spaces for the transition task in study 2: (a) office, (b) kitchen, and (c) living room.

5.14 (a) The average time it took for participants to complete all four scenarios for an interface condition (in seconds); (b) the average head rotation (in degrees/second); and (c) the average System Usability Scale (SUS) score (±S.E.).
5.15 (a) The average ratings for the three questions regarding Agency; (b) the perceived accuracy level of the automated apps (±S.E.); and (c) the overall preference rankings of the three interfaces. .................................................. 123

6.1 Using Glanceable AR interface in three everyday scenarios: (a) working in front of a desktop computer with glanceable widgets residing at the edge of the physical monitor; (b) cooking with recipe and timer widgets following the user for hands-free access of information; (c) walking outside with music, fitness and map widgets following the user; a notification notifies the user about accomplishing the daily step goal. ............................................................... 139

6.2 Screenshots of the video prototype: (a) working in front of a desktop computer with glanceable widgets residing at the edge of the physical monitor; (b) cooking with recipe and timer widgets following the user for hands-free access of information; (c) walking outside with music, fitness and weather widgets following the user (Video prototype available here: https://youtu.be/qsDVDP7wA-E) 140

6.3 (a) UEQ results for both overall and separate features shown in the video prototype; (b) ratings of ease of access and distraction for six tasks (higher ratings mean easier access and lower distraction) (±S.E.). ........................................ 148

6.4 Gaze-contingent interaction with dwell: (a) user starts to gaze at the calendar; (b) user keeps gazing at the calendar with the progress circle filling up; (c) after one second, the calendar expands with more detail. ................................. 154

6.5 An illustration of the Fixation Glance (FG) technique: (a) widgets are represented as small targets to avoid occluding users’ view when the user is looking at the real-world environment behind the target; (b) when the user converges their gaze at the depth of the target, the widget expands and appears. .... 155
6.6  (a) Using raycasting to select and move widgets; (b) Adjusting news category and notification levels in the main menu. ................................. 155

6.7  (a) UEQ results for study 2; (b) average frequency of glancing at each widget, red indicates a quick glance less than a second and blue indicates a long glance over one second to expand the widgets; (c) percentage of time participants spent in each mode on average for each session; (d) duration in seconds participants spent gazing at each widget on average for each session (±S.E.). 159

6.8  (a-c) The new designs of the AR applications with added interactivity; (d-e) the new menu design with hand trigger and gaze interactions (i.e., Dwell, Blink) for selections. .................................................. 167

6.9  (a) The new design structure for the Gmail, Calendar, and News applications. Two arrows appear on top of the digital apps for browsing through multiple pages. Gazing at an interactive item (e.g., an email subject) would enlarge it, making it easier to glance at. (b) After reading the email, the user could glance at either the Trash or the Star buttons to perform the actions correspondingly; (c) Each app has a built-in menu with three features: follow (left), mute (middle) and close (bottom). .............................................. 168

6.10 (a) The new design of the Follow feature. An icon replacing the dot to enhance awareness of the minimized applications; (b) A wiggling animation was used to indicate successful detection of user blinks; (c) After two consecutive blinks, the app was activated. .......................................................... 170

6.11 The average UEQ scores and the benchmark results. ............................ 172

6.12 The average time participants spent on viewing/interacting with each of the AR apps. ............................................................. 173
6.13 (a) The percentage of time in which users had 1/2/3/4 apps following them; (b) The percentage of time in which the users had 7/6/5/4 apps opened (the percentage of time that the users closed more than 3 apps is below 0.2%); (c) The percentage of time in which the users enabled 6/7 apps notification (the percentage of time that the users muted more than 1 app is below 0.03%).

6.14 (a) The percentage of time in which users used Blink vs. Dwell as a confirmatory input; (b) percentage of time users used each interactions for glancing at the AR widgets during mobile use cases.

7.1 If user A is reading a piece of virtual content, Glanceable AR interfaces could vary in translucence levels that visualize user A’s interactions to user B to achieve more efficient and comfortable collaborations (the blue diamond indicates that the user is currently looking at a piece of digital information rather than the physical world).


<table>
<thead>
<tr>
<th>Table Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Characteristics of the three interfaces in terms of ease of access, awareness of information, and unobtrusiveness of display.</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>The five techniques to activate or deactivate virtual information and the hypothesized trade-offs.</td>
<td>63</td>
</tr>
<tr>
<td>5.1</td>
<td>Hypothesized contextual preferences on interfaces.</td>
<td>113</td>
</tr>
<tr>
<td>5.2</td>
<td>Characteristics of the three physical rooms used in the study and the list of automated applications in the <em>Some</em> interface.</td>
<td>116</td>
</tr>
<tr>
<td>5.3</td>
<td>The number of votes each interface received under each contextual situation, with the interface condition with the most number of votes bolded.</td>
<td>125</td>
</tr>
<tr>
<td>6.1</td>
<td>Most frequently occurring codes and their frequencies in four categories: benefits of the interface, drawbacks of the interface, wish-to-have features, and other scenarios in which the interface could be useful.</td>
<td>149</td>
</tr>
<tr>
<td>6.2</td>
<td>Interactive features added to selected applications.</td>
<td>168</td>
</tr>
</tbody>
</table>
List of Terms

AR (Augmented Reality) - Real time combination of real objects and virtual objects registered in three-dimensional space that is interactive [10].

Eye Tracking - A technology for real-time measurements of eye positions, directions, and movements.

FoV (Field of View) - The extent of observable world through the display, usually measured in visual angle.

HCI (Human-Computer Interaction) - Research domain focusing on the design and the use of computer technology, and the interfaces between users and computers [1].

HWD (Head-Worn Displays) - A display device, worn on the head that has a small display optic in front of one (monocular) or each eye (binocular).

Input Modality - A modality is the classification of a single independent channel of sensory input/output between a computer and a human [86].

Interaction Technique - A method allowing a user to accomplish the task via the user interface (UI) [93].

LoD (Level of Detail) - The amount of information presented in a piece of graphical representation.

MR (Mixed Reality) - A set of approaches, including both virtual reality (VR) and augmented reality, in which real and virtual information is mixed in different combinations along the reality-virtuality continuum [113].
UI (User Interface) - The medium through which the communication between users and computers take place.

VE (Virtual Environment) - A synthetic, spatial world seen from a first-person point of view. The view is under real-time user control [93].

VR (Virtual Reality) - An approach that uses displays, tracking, and other technologies to immerse the user in a virtual environment [93].
Chapter 1

Introduction

1.1 Motivations

People encounter a variety of information needs in their daily lives. Information could be needed anywhere, anytime to assist in decision-making and execution of certain tasks [33, 44]. For example, people would want to check their to-do lists to determine schedule for today, or look up the weather information to decide the clothes to wear. Research has found that such information needs are often triggered in mobile settings when users move to different locations and need information on-the-go [32, 144].

Mobile computing devices have been developed to support users’ ubiquitous needs of information. Mobile phones, as the most pervasive personal computing device nowadays, provide convenient ways to obtain information. However, its handheld nature and touch interfaces could introduce limitations in some scenarios [144]. 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 to perform. Wearables, especially smartwatches, are capable of displaying information, such as the next calendar event, to users without the need for most the steps above. However, the small screen size limits the amount of information being presented.
at any time. As such, in everyday scenarios, it happens frequently that information needs are postponed or left unaddressed by the users due to the workload required to access the information, potential disruption to the tasks the users are engaging with, and the limited cognitive bandwidth available to perform the inquiries [33, 144].

Augmented Reality (AR) head-worn displays (HWDs) have been long considered as the next-generation personal computing device. They have the potential to support easy and fast access to 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 [56].” In 2010, Want wrote that “augmented reality could become an indispensable tool of the future in much the same way we have come to reply on the cell phone today [156].” Conventional AR systems have typically been expensive, bulky, and technologically limited. Thus, they have been used primarily for special-purpose tasks, such as navigation, training, or maintenance. With recent advancements in display technology and tracking, however, AR systems are becoming increasingly accessible and affordable. It is now possible to deliver convincing AR experiences to users via relatively lightweight AR HWDs. For example, HWDs such as the Spectacles by Snap Inc\(^1\), the ThinkReality A3 Smart Glasses by Lenovo\(^2\), and the nreal Light\(^3\) have form factors close to a pair of eyeglasses and are designed with a vision to support general-purpose everyday use cases. These devices may eventually replace our smartphones, enabling continuous, universal, and pervasive augmentations in the real-world to assist people with their everyday tasks anywhere and anytime [13, 65].

However, for AR HWDs to be widely applicable for everyday uses, research is necessary not only about how to improve the optics, tracking, and form factor of the display, but also about how information needs to be delivered to users, perceived by the users, and

\(^1\)https://www.spectacles.com/
\(^3\)https://www.nreal.ai/light/
interacted with by the users in everyday contexts. There are two major differences between AR HWDs and traditional mobile devices. First, AR does not constrain content displays on 2D screens with limited sizes, but allows content display on an infinitely large 3D canvas (i.e., the real-world environment). Positioning of digital information is no longer limited to the physical displays. Instead, information could be body-referenced, device-referenced or world-referenced depending on the scenario of use. Second, AR HWDs are worn on the head and always-visible to the eyes, which “forces” the users to see the displayed content (as long as it is in the FoV); users cannot look away from the display. With augmented virtual information becoming pervasive and always available, virtual content could occlude the objects of interest in the real-world, distract or disrupt our tasks at hand, become overwhelming and cause information overload. Meanwhile, everyday AR content needs to be interactable to provide users with the controllability to access the right service at the right time. Such interactions, if not designed carefully, could require cognitive and physical effort, making them impractical or even dangerous to use in already-demanding everyday situations. As pointed out by Grubert et al., we need to design “appropriate information display and interaction, which is unobtrusive, not distracting, and is relevant and safe to use [65],” which remain an unaddressed yet critical challenge for future always-on AR HWDs.

1.2 Research Questions

In this research, we aim at enabling a pervasive yet unobtrusive AR future by addressing the following research questions (RQ):

RQ1 - How can we support non-distracting yet efficient and natural information acquisition with AR HWDs? To make AR HWDs versatile to use in everyday scenarios, we have to research information display strategies that are non-distracting, but efficient
and natural to use. The goal is for users to be able to focus on their tasks at hand in the real-world while accessing required information without dividing their attention, switching context regularly, or being disturbed by the AR content.

**RQ2** - **How can we prevent AR content from being obtrusive and occluding the real-world objects of interest to the users?** Future always-on AR glasses could support easy access to digital information by displaying it directly in a user’s FoV. However, with augmented virtual information becoming pervasive and always-available, information could become noticeable and occlude real-world objects that are of interest to the users. How can we resolve unwanted occlusion of the real-world environment by AR content without compromising its accessibility, thus enabling efficient acquisitions of both the AR content and the real-world environment?

**RQ3** - **How can we support relevant and scalable acquisition of AR content during spatial transitions?** In commercial AR systems, AR content often stays world-referenced and remains at a fixed location until it is manually moved by the users. However, in real-world situations, users often transit spatially from one location to another while still needing to access information. For example, users might move from the living room to the kitchen while needing to know what to buy later in the grocery store, or move from the office to the living room while wanting to know the weather of today. With existing UI solutions, users would lose access to the AR applications placed in the previous location after transitions, thus having to manually reopen them from a menu interface and place these applications in the new space. Such interactions could be tedious and cumbersome to perform repeatedly. How can we design AR interfaces which are transition-friendly and scalable, so that users can seamlessly bring AR applications with them from place to place for information access on-the-go?

**RQ4** - **What are user perceptions and acceptance towards using pervasive, always-
on AR systems in authentic everyday contexts? Controlled experiments usually happen in a controlled lab space with predefined scenarios and tasks. However, in actual everyday scenarios, interactions with AR HWDs: (1) would be unsupervised without the presence of the experimenter; (2) would be informal and unstructured, happen anywhere and anytime without a clear beginning or end; (3) could be interrupted by random uncontrolled events; and (4) could involve operations of multiple concurrent activities [4]. Therefore, it is necessary to deploy our AR interfaces with real-world users in authentic everyday-life contexts to understand genuine user perceptions and needs, which often cannot be reflected in controlled evaluations in a lab space. As such, the fourth research question of this dissertation is about evaluating proof-of-concept AR systems to support people’s information needs in real-world settings to understand user perceptions and acceptance in authentic everyday contexts.

1.3 Methods and Scope

1.3.1 The Glanceable AR Paradigm

With the goal to address all RQs, in this dissertation, we propose Glanceable AR, an interaction paradigm that allows efficient, natural, and non-distracting information display and acquisition with AR HWDs. Inspired by previous work in ubiquitous computing and calm technology by Weiser and Brown, we believe that future computing services will be everywhere, should recede into the background of our lives, should inform but not demand our attention, and should be easily available when we need them [106, 159, 162]. In Glanceable AR, virtual information resides in the visual or display periphery to stay unobtrusive, but can be prioritized and accessed through a glance when needed (see Figure 1.1). Our definition of Glanceable AR, although was restricted in the initial concept demonstrated in chapter 3,
is broad and refer to AR information display strategies in which:

- AR content is secondary, non-invasive, and non-distracting. It stays in the visual or display periphery, and is readily available to support users’ tasks.

- AR content draws minimal attention from the users when not needed, and it is up to the users when, where, and how to prioritize it.

- The information displayed in the AR content could be quickly accessed and understood through glances of the eyes, backed up by lightweight interactions.

- AR content has high mobility, react to the users’ movements in the real world for on-the-go access regardless of physical locations.

In this dissertation work, our research centers and revolves around the Glanceable AR concept by iteratively designing, refining, evaluating, and deploying Glanceable AR interfaces, both in controlled and real-life everyday contexts (see Figure 1.2). We focused on four aspects of Glanceable AR, each addressing one of the four RQs mentioned above, including mitigating distractions (see chapter 3), resolving occlusions (see chapter 4), supporting scalable transitions (see chapter 5), and real-world evaluations (see chapter 6).

### 1.3.2 RQ1: Mitigating Distractions

To keep virtual content from distracting the users in AR HWDs, we designed a set of glanceable interfaces with the goal of allowing users to access information easily without their primary tasks being disrupted in the real-world. We then evaluated them in multiple dual-task scenarios, with primary tasks of different workload and secondary tasks requiring different levels of attention on the virtual content in AR. Please refer to chapter 3 for project details.
1.3. Methods and Scope

1.3.3 RQ2: Resolving Occlusions

There are multiple ways of resolving the occlusion issues of AR content on the real-world environment. Previous work in view management explored adapting the layout, position, and size of virtual annotations to enable a more clear and organized view [14]. In chapter 4, we look into a unique way of minimizing occlusions of virtual information display. In our approach, AR content was displayed as low LoD, semi-transparent icons by default, but could be activated to higher LoD only when the users wanted to know more about it. We first conducted a study to validate the occlusion issues in AR HWDs during a visual search task. Then we proposed a design space, which led to five interaction techniques to activate virtual content with gaze, hand, and head-based interactions. We evaluated them in both seated and walking contexts.

1.3.4 RQ3: Supporting Scalable Transitions

In chapter 5, we focus on designing interfaces which support information needs during user transitions across multiple locations in the physical world. Specifically, we explored the
Figure 1.2: An illustration of the structure and organization of this dissertation work.

The trade-offs among three conditions: (1) None: the system does not bring any apps with the users, so they always have to manually reopen the app through the menu; (2) All: the system automatically brings all the apps users left in the previous space to the new space, no matter whether they are needed by the users or not; (3) Some: the system brings a subset of apps which are of high relevance to users’ transition destination, and users can make changes to the recommended list. We first evaluated the three conditions in a controlled scenario with different total number of virtual apps to model user performance. Then, we conducted an in-situ transition study, in which users were asked to perform embodied tasks in the physical world while needing access to virtual information on-the-go. We evaluated the usability of the three interface solutions while taking into account a set of contextual factors, including
priority of the real-world environment, task specificity, and the workload of the task that user engages with in the real-world environment.

1.3.5 RQ4: Real-World Evaluations

After addressing these common challenges in information displays in AR HWDs, to address RQ4, we deployed our systems to real-world contexts to collect data about usage, perception, and acceptance. In order to capture genuine user perceptions and needs, we utilized three approaches: (1) a video prototype, evaluated through an online survey to reach a broad audience with different ages, backgrounds, and expertise; (2) real-world deployments of proof-of-concept Glanceable AR systems with personalized information display; and (3) a real-world deployments of a proof-of-concept feature-rich Glanceable AR system with expert evaluations. Throughout the dissertation work, we developed two proof-of-concept Glanceable AR prototypes. Both prototypes were linked with users’ personal data, including email, calendar, to-do list, fitness, as well as information in the physical world which was of high relevance to users’ lives, including local time, weather forecast, and news information. As such, the prototype became an alternative to mobile phones for users to rely on to assist their everyday tasks. In two studies, we deployed our prototype to real-world users, with and without prior experience to AR, and obtain quality feedback about the user experiences of using Glanceable AR in the wild. Users were asked to use the system for multiple sessions a few consecutive days concurrent to their everyday lives covering all sorts of scenarios, while maintaining a diary documenting their thoughts after each session of use. We obtained quantitative and qualitative results that reflect genuine user perceptions and requirements while using AR HWDs as alternatives to mobile devices in everyday lives. Please refer to chapter 6 for more details.
1.4 Contributions

Our work contributes to the pervasive adoption of future always-on AR HWDs in the following ways: (1) Identification of challenges when using AR HWDs in everyday scenarios for information access (chapter 6); (2) Proposing interface solutions that address the distraction, occlusion, and transition issues of AR content in common everyday scenarios (chapter 3, chapter 4, chapter 5, and chapter 6); (3) Evaluations of these interface solutions in controlled laboratory studies to reveal the user experience and trade-offs among them (chapter 3, chapter 4, and chapter 5); (4) Implementations of proof-of-concept Glanceable AR systems with state-of-the-art AR HWDs (chapter 6); (5) Deployments of these systems in authentic everyday use cases to understand user perceptions and requirements (chapter 6).
Chapter 2

Review of Literature

2.1 The Evolving Eras of Computing

Decades ago, the invention of computers marked the opening of the first wave of computing, *mainframes*. Computers were huge, expensive and power consuming, so that many people had to share a single one. As technology advanced, computers became more lightweight and powerful. People were able to carry around their own personal computers to assist their everyday activities and executions of tasks. Here came the second wave of computing, *personal computing*. People started spending many hours per week looking at computer screens or televisions. While computers were becoming increasingly accessible and pervasive, they seemed to draw people’s attention away from the physical world that they live in.

In 1988, Mark Weiser coined the third era of computing while he was the director of the Computer Science Laboratory in Xerox PARC, the era of *ubiquitous computing*, in which he envisioned that in the near future, a single user would be interacting with multiple computers simultaneously [161]. Figure 2.1 shows a plot of number of devices over time in the three eras of computing [156]. In Weiser’s article *The Computer for the 21st Century*, he wrote a well-known quote: “The most profound technologies are those that disappear [161].” He believes that in the future, computers and processors will become invisible and recede into the background of our lives. They seamlessly coordinate together and react to users in a variety of everyday life/work context, and maybe users will not even notice the involvements of these
computers. The term *embodied virtuality* refers to the procedures of drawing computers out of their electronic shells and fading into the invisibility. Weiser also referred to this as *calm technology* or *invisible computing* [160, 162, 163]. By de-emphasizing the existence of computers, ubiquitous computing (ubicomp) pushes the vision of computer interfaces back to the physical world around the users [158].

Augmented Reality is one of the few technologies that is designed to emphasize the real world. Different from virtual reality, which immerses users in a computer-generated virtual world, AR embeds virtual content in the real-world environment [10]. The real world becomes the major carrier of information, which has to be perceived to comprehend the virtual information that comes with it. Liberati commented that both AR and ubicomp are technologies that try to “escape the cyber vacuum of VR”, and they try to “bring the computing power to our world” [97]. Recent AR displays are becoming more and more lightweight and powerful. They have the potential to fulfill the visions of ubiquitous computing by offering pervasive, non-intrusive, and highly mobile information displays within the real-world environment.
2.2 Pervasive and Everyday Augmented Reality

Conventional AR systems have typically been expensive, bulky, and technologically limited. Thus, they have been used primarily for special-purpose tasks, such as navigation, training, or maintenance. With recent advancements in display technology and tracking, however, AR systems are becoming increasingly accessible and affordable. It is now possible to deliver convincing AR experiences to users via relatively lightweight HWDs. Shortly, AR content could become everywhere and always-available, augmenting both the physical world and people’s cognition by providing easy and quick access to a variety of information and services.

In 2016, Grubert et al. proposed “Pervasive AR,” which refers to “a continuous, omnipresent, and universal augmented interface to information in the physical world” [65]. Different from conventional AR experiences which have been mostly special-purpose, prototypical, obtrusive and sporadic, pervasive AR interfaces set their visions at prolonged and continuous uses. For example, an everyday AR user could view calendar information (see Figure 2.2 (a)), check reviews of restaurants, and conduct productivity work directly through the glasses without requiring access to any physical display. However, with augmented virtual information being pervasive and always available, certain issues could be introduced in everyday scenarios. The virtual information might become overwhelming, occlude real-world objects of importance, distract or disrupt users from their real-world tasks (see Figure 2.2 (b)), and further lead to privacy issues when interactions are visible by co-present others. Such provocative and dystopian future visions of AR have been increasingly concerning in society, in which people are worried about virtual information taking over the real world, delivering irrelevant content, and consuming the majority of their cognitive bandwidth [2, 3]. As such, if AR experiences are to be omnipresent and pervasive, they have to be tailored, individualized, adaptive and unobtrusive. As pointed out by Grubert et al., when AR experiences become omnipresent,
we need to design “appropriate information display and interaction, which is unobtrusive, not distracting, and is relevant and safe to use [65],” which shares similar visions with ubiquitous, pervasive, invisible computing and calm technology.

However, there has been lack of research regarding how information could be displayed in AR without distracting the users, while being easily accessible and understandable. In this dissertation, we aim at addressing these challenges. Specifically, we looked into unobtrusive ways that information could be displayed (see chapter 3), acquired (see chapter 4), moved (see chapter 5), and used for general-purpose everyday scenarios (see chapter 6) in AR HWDs in the context of everyday uses.

### 2.3 General-Purpose Information Displays in AR

The majority of existing AR applications are designed for special-purpose use cases such as training simulations [60] and knowledge work [66, 96]. However, AR systems, especially AR HWDs, have long been envisioned as general-purpose computers to assist people’s everyday
tasks. Back in 2002, Feiner said that future AR HWDs will become “much like telephones and PCs,” and display information “that we expect to see both at work and at play” [56]. With the advancements in AR systems in recent years, we are coming closer to this envisioned future. Google Glass was an early attempt to apply wearable HWDs to general-purpose everyday use cases [145]. Recently, more smart AR glasses have been announced. For example, the nreal Light glasses come with built-in 6 Degree of Freedom (DoF) tracking and hand tracking capabilities, with a form factor similar to a pair of conventional eyeglasses.

Current commercial HWDs to date for general-purpose use can be categorized into two types: (1) untracked wearable displays in which information is fixed to the display (e.g., Google Glass¹); and (2) wearable displays with tracking capabilities to register information to the physical environment (e.g., nreal Light², Snap Spectacles³). The benefit of both types is that they are worn on the head and directly visible to the eyes, which makes rapid information access viable. However, an advantage that only a tracked display can offer is that it allows registration of virtual content to the real-world (locations, objects, or even users’ body) instead of only on the display.

With tracked displays, virtual information can be placed freely anywhere in the 3D space. However, placing it near a real object in the physical world as a frame of reference could make it easier to perceive and interpret by the users [138]. Strategies for virtual information display in the literature can be categorized as world-fixed, object-fixed, head-fixed, body-fixed or device-fixed [21, 93]. In AR specifically, world-fixed, display-fixed and body-fixed are popular layouts to register virtual windows. An early prototype by Feiner et al. in 1993 described the three strategies: (1) fixed relative to the HMD; (2) fixed relative to real-world locations and objects; and (3) fixed to a sphere surrounding the users [55]. ARWin was a desktop

¹https://www.google.com/glass/start/
²https://www.nreal.ai/light/
³https://www.spectacles.com
workspace augmented by virtual calendar, weather and web browsers [45]. This information was world-fixed to physical surfaces around the desktop. Starner and Billinghurst proposed a spatial display metaphor, in which virtual windows were body-fixed to a cylinder surrounding the user [16]. They found the approach was faster in locating information than presenting all information at one position. Ens et al. proposed personal cockpit, an system with adaptive displayed-fixed, body-fixed and world-fixed layouts for everyday mobile applications, which resulted in 40% improved task-switching time as compared to display-fixed layouts [52]. Lages and Bowman evaluated an adaptive walking interface for AR HWDs. Their interface could switch between world-fixed and body-fixed layouts based on user input. Their result emphasized that it is important to prevent virtual information from blocking real-world objects or environment that is of interest to the users.

Existing work in view management has explored meaningful ways of placing virtual content on the view plane to avoid occluding each other and ensure a structured view [14]. However, there has been a lack of research regarding how virtual information should be displayed in non-intrusive but consumable manners, especially in mobile scenarios. One way to alleviate the distraction and occlusion issues caused by augmented virtual information is adaptive information display. In 2003, Vertegaal first proposed Attentive User Interfaces (AUI), a system that was continuously aware of what users pay attention to and adapted its services [153]. In 2019, Lages and Bowman explored an adaptive walking UI, in which AR windows were placed adaptively around users’ body or on the wall based on user input [91]. Lindlbauer et al. explored automated placements of AR content based on task and eye-tracking data [98]. Cheng et al. explored automatic adaptation of UI’s spatial layouts based on environmental changes when users move to different locations [30]. Their results shed light on the potential of AR systems to predict user needs and assist the placements of AR UIs. Another way of displaying the information adaptively is through users’ explicit interactions [47, 69, 125, 134].
Such methods put more efforts on the user side as compared to automatic detection on the system-level, but could be more predictable and accurate, especially in everyday situations when contextual information changes frequently.

In this dissertation work, we explored different ways of displaying virtual content to allow easy consumption of information while avoiding being distracting or obtrusive to tasks in the physical world (see chapter 3 & chapter 4). We cover a variety of scenarios and tasks, including stationary and mobile scenarios with lightweight and heavy tasks.

2.4 Peripheral Vision, Display, and Awareness

In information design, there are multiple strategies for keeping users aware of important information. The peripheral awareness strategy places information into users’ peripheral attention in a way that it is easily accessible without being distracting [25]. In the field of visual perception, periphery is defined as the field of vision outside the corner of our eyes, which helps us to view information that lies outside the central vision [84]. Weiser and Brown defined the “periphery” as “what we are attuned to without attending to explicitly” [162, 163]. In this dissertation work, we adopted a combination of both definitions. We define “periphery” as the field of vision outside the central vision that we are not paying direct attention to normally, but could be accessed through a glance of eyes. Glanceable AR sets to utilize users’ periphery so that the users are empowered to take control of the content and initiate interaction based on their needs[109, 162]. 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 [162].

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
Embedding information in the periphery can give users opportunities to learn more, do a better job, or keep track of less important tasks [107].

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 [162]. 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 users’ primary tasks, and allow quick intake of information through a glance at the periphery [108]. InfoCanvas is a peripheral display that allows communication of quantitative information through visual representations of known objects on a secondary display [114]. Sideshow and Scope are peripheral displays that sit at the border of the screen showing general information such as incoming emails and tasks [25, 151]. ShutEye is a peripheral display that encourage awareness of healthy sleep behaviors [12]. Ambient displays are a specific type of peripheral displays that utilize physical materials or lifelike forms distributed in the environment to convey information [71, 75, 164]. These projects emphasized the importance of information awareness, and gauged the viability of obtaining peripheral information in a controlled space. In this dissertation work, our idea centers around the utilization of “periphery” as the carrier of information to enable quick “glances” of information with low cognitive effort in practical multi-tasking setups. Peripheral vision is defined as the area further away than 30 degrees to central view (further categorized as near: below 30 deg; mid: 30 deg – 60 deg; and far: 60 deg – 110 deg). However, due to the limited FoV of current AR displays, it is hard to display AR content at the far periphery. For the scope of this work, we define periphery as both the edge of the AR display FoV and the periphery of the human vision.
2.5 Input Modalities in AR/VR

Existing research has explored an extensive set of input methods for interacting with virtual content in AR HWDs, each with its pros and cons and appropriateness for different contexts of use. Among them, controller, head, hand, and gaze-based interactions are the most promising ones given the tracking capability of current AR HWDs [105, 138, 142, 167].

In existing work, Davari et al. explored pointing and clicking-based interactions with a hand-held controller [42]. However, controllers are less likely to be always available in daily situations. Hands, as a natural "controller" that comes with our bodies, are gaining popularity with the advancements in computer vision. For example, Xu et al. explored using hand gestures for word selections in text entry [167]. Mutasim et al. compared hand-based pinch and dwell gestures for target selections [117]. Gaze and head-based interactions have been applied when users’ hands are unavailable for some tasks. Gaze, as an intuitive indication of the direction the user is looking at, has been primarily used to trigger gaze-contingent adaptive information displays. Pfeuffer et al. explored adapting the transparency and information level in the virtual content through gaze directions [125]. Head-based interactions have been explored as a potential input method as well. For example, Diverdi et al. proposed LoD interfaces, in which virtual content was displayed with different levels of detail (LoDs) based on the distance between the camera to the virtual content [47]. In recent work, Yu et al. explored target selection with head movement in the depth dimension [170]. Shi et al. explored mode-switching with head gestures only [142]. This work demonstrates the viability of different input modalities in special-purpose use cases.

In this dissertation work, we explored various input methods to trigger adaptive behaviors of virtual information displays in different everyday contexts (see chapter 3 & chapter 4). Specifically, we focused on adaptive frame of reference, adaptive LoD, and adaptive interac-
2.6 Evaluations of promising everyday AR systems

Envisionments of promising technologies are common in the HCI community. Understanding acceptance, preferences and requirements of intended users are crucial prior to actual implementations to save cost and effort. Previous research has mainly applied three methodologies: (1) Distribution of online surveys to obtain feedback from a broad population. For example, Popovici and Vatavu obtained feedback from 172 participants on interactions and scenarios for an envisioned AR television \[127\]. (2) Employment of visual prototypes to convey design ideas. In the same study, Popovici and Vatavu described use cases of AR television via video prototypes and illustrations \[127\]. Häkkilä et al. conveyed the concept of AR windows for cars via video see-through displays \[67\]. (3) Field studies to understand real-world user acceptance and requirements.

The term “in the wild” in HCI research refers to performing studies with users in uncontrolled environments rather than in the laboratory \[23\]. Steed et al. proved the viability of in-the-wild VR experiments in a study that explored embodiment and presence with unsupervised users in commercial VR headsets \[146\]. Field studies, as a methodology of conducting in-the-wild research, collect data in users’ contexts. Compared to traditional user studies that take place in a lab with structured and artificial tasks, field studies can be more powerful to “understand the end user’s natural behavior in the context of his or her everyday environment \[39\].” Existing research proved the viability of field studies to obtain quality feedback from users in real-world tasks and environments with AR systems \[89, 92, 116, 132\]. For example, Ventä-Olkkonen et al. explored the concept of a AR city via field studies and conceptual prototypes \[152\]. This research shows field studies as an approach to reveal real-world use...
cases for non-existing technologies. In this dissertation work, we employ similar approaches to distribute and evaluate our idea of using Glanceable AR for everyday scenarios and tasks by deploying it with everyday users longitudinally in the real-world (see chapter 6). Our results contribute to a better understanding of user experience, perceptions, and acceptance of using AR to assist their tasks in unstructured everyday activities.
Chapter 3

Designing for Unobtrusive and Non-Distracting Information Access

3.1 Overview

In this chapter, we cover our solution to displaying information in AR HWDs without being obtrusive or distracting to users’ primary tasks in the real world. 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 [91]. 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 can efficiently and naturally access the information we need in AR HWDs.

In this research, we propose *Glanceable AR*, an interaction paradigm for accessing informa-
tion in AR HWDs. In Glanceable 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 users’ 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.

This work has been published at IEEE VR 2020.

3.2.1 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 [108]. 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 [109]. 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.

3.2.2 Initial Design Principles of Glanceable AR

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 drafted four initial design principles for our Glanceable AR interfaces backed up by literature in AR/VR, spatial memory, and ubiquitous computing.

**DP1.** Virtual content should be fixed to the user.

---

1 Please refer to subsection 1.3.1 for the refined definition.
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.

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 users’ body [55]. 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 [91]. 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 means that instead of strictly world-fixed layouts, future information displays should be fluid and fixed to the users. As such, utilizing display-fixed and surround-fixed layouts, or designing strategies to allow AR content to transition between multiple coordinate systems, may be more preferable [15, 52].

DP2. Spatial Distribution. Billinghurst and Starner proposed two metaphors of accessing information in AR HWDs [16]. 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 (i.e., a heads-up display (HUD)). 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, allocate one piece of content to each direction with restrictions to four applications at a time. 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.

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 [133]. 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 [115]. 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., [49, 110]). In near-eye displays specifically, eye-tracking has been proven to be faster compared to finger-pointing [148]. 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.

This principle may be ideal for future AR systems, but is loosened for the scope of this dissertation. Hands/controllers are necessary to make precise inputs/manipulations due to limitations in current interfaces/technology.
Figure 3.1: Three Glanceable AR interfaces proposed in this section: (a): Eye-Glance (EG) interface (HUD), in which virtual content resides at the edge of FoV and is fixed to users’ head; (b) Head-Glance (HG) 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 (GS interface, in which virtual content can be summoned into FoV by gazing at the periphery for 0.5 second.).

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 [17]. 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.
3.2.3 Glanceable AR Interfaces

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

Eye-Glance (EG) Interface

The most basic realization of the Glanceable AR approach is a simple Heads-Up Display, which we call an “eye-glance” interface (see Figure 3.1a). 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 users’ 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\(^3\) and Focal smart glasses\(^4\)) 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. Figure 3.2 (a) shows an implementation of the eye-glance interface in AR.

**Head-Glance (HG) Interface**

To make content more unobtrusive, the head-glance interface places content outside users’ forward field of view and fixes the content to users’ body rather than the display. To access content, users simply turn their heads towards one side of the periphery (see Figure 3.1 (b)). 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. Figure 3.2 (b) illustrates an AR implementation of the head-glance interface.

**Gaze-Summon (GS) 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 users’ gaze direction [62, 131]. Based on our review, gaze-contingent

\(^3\)https://www.google.com/glass/start/
\(^4\)https://www.bynorth.com/focals
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 Figure 3.1 (c)). 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”[82], we use a dwell technique [81]. 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. Figure 3.2c 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 rather than to the real-world [10], 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 [55, 85].

Trade-offs Among Interfaces

Table 3.1 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
Table 3.1: Characteristics of the three interfaces in terms of ease of access, awareness of information, and unobtrusiveness of display.

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Accessibility</th>
<th>Awareness</th>
<th>Unobtrusiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye-Glance</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Head-Glance</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Gaze-Summon</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
</tr>
</tbody>
</table>

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.
3.3 Evaluations of Glanceable AR in Walking Scenarios

3.3.1 Research Questions

We aim at answering two research questions in this study:

- **RQ1.A.** How distracting or intrusive Glanceable AR interfaces would be towards users’ primary tasks in the real-world while used for information access?

- **RQ2.B.** What are the user experience, workload, and performance of using Glanceable AR interfaces when different levels of attention are required on the AR content?

3.3.2 Hypothesis

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.

### 3.3.3 Participants

We recruited 18 participants (4F/14M) 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).

### 3.3.4 Tasks

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 trade-offs 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.

We used a $3 \times 2$ 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.

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. Figure 3.3 illustrates the task in a top-down view.

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 Figure 3.2a) 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 head-
3.3. Evaluations of Glanceable AR in Walking Scenarios

![Figure 3.3: 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.](image)

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 10.8 degrees horizontal and 3.4 degrees vertical on the display. Questions were asked verbally by the system randomly about information in one of the four pieces of virtual content; 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 [144]. Participants had to search for the corresponding virtual content, check its information, and answer the questions verbally. 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 Figure 3.3d), 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.
3.3. Evaluations of Glanceable AR in Walking Scenarios

3.3.5 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 \[ F_1 + C_1 + 2(F_2 + C_2) + 4(F_3 + C_3) \], 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 Figure 3.3 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 raw NASA TLX workload questionnaires to gauge the usability and workload of the interfaces [22, 70]. 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.


3.3.6 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.
3.3.7 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 pairwise 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 chapter.

Primary Task Performance

Figure 3.4 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.

Secondary Task Performance

In the discretionary task, we collected 12 (number of questions per condition) × 18 (number of participants) × 3 (number of interfaces) 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. Figure 3.5 (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 (Figure 3.5 (right)), we again averaged the time for each
3.3. Evaluations of Glanceable AR in Walking Scenarios

Figure 3.5: Bar charts of time (in seconds) taken to answer question/report lead changes in discretionary (left) and monitoring (right) conditions. (± S.E.)

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: $p = .210$).

NASA TLX Workload & SUS Score

Figure 3.6 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 (Figure 3.7), 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).
3.3. Evaluations of Glanceable AR in Walking Scenarios

Figure 3.7: Mean SUS score categorized by interfaces for discretionary (LEFT) and monitoring (RIGHT) tasks (±S.E.)

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%).

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].

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).
3.3. Evaluations of Glanceable AR in Walking Scenarios

Figure 3.8: Mean occurrences of (Left) looking around to find virtual human; (Right) adjusting distance with virtual human after a question was asked in discretionary task (±S.E.)

Figure 3.8 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.

3.3.8 Summary of Findings

We summarize our findings as follows:

- EG resulted in the best primary walking task performance in both discretionary and monitoring conditions.

- HG resulted in the worst primary walking task performance, but participants were able to achieve a similar level of secondary task performance in terms of discretionary acquisition and monitoring of virtual content.

- GS resulted in the worst secondary task performance in both discretionary and moni-
EG received the highest usability ratings, and was the most favored by participants when continuous attention was required on the virtual content.

### 3.3.9 Discussions

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
3.4. Conclusions

In this chapter, 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.
Chapter 4

Resolving Occlusion Issues of AR Content

4.1 Overview

Recent research highlighted an important issue to be addressed for everyday AR interfaces, which is how to handle the potential occlusion of virtual information overlaying the real-world environments that are of interest to the users [42, 91]. In chapter 3, we proposed Glanceable AR, in which virtual information resides at the peripheral vision to avoid occluding users’ view when not needed [100, 102]. Although Glanceable AR is a promising solution to enable always-on information display, whether it completely addresses the occlusion challenge remains questionable. Virtual information in the periphery can still block users’ view when they intend to glance at the real world behind the virtual content. To make glanceable information more unobtrusive, an adaptive information display strategy can be applied, in which virtual content is minimized when not needed, but can be activated through interaction techniques to show more detailed information [47, 98, 111, 125, 126]. In this way, the real world is “prioritized”, and users can explicitly control what they see depending on their needs [42, 125].

However, effective activation techniques for glanceable content must address several challenges. First, the techniques have to be rapid, as users need to be capable of performing
them within the duration of a quick glance. Second, they need to be lightweight, so that users can allocate their limited cognitive and attentional resources to their primary tasks and to events in the environment, rather than diverting cognitive effort to the interface. Third, they must be reliable, since unintentional activation could occlude important information, and failure to activate on demand could lead to distraction and frustration.

In this research, we first conducted a user study to validate the occlusion and distraction issues in Glanceable AR interfaces. We then addressed the design challenges by proposing a set of design considerations for developing techniques to explicitly activate virtual content in Glanceable AR interfaces. Next, we proposed five techniques for activating the minimized glanceable widgets, and evaluated them in controlled experiments. Two common everyday scenarios were replicated in the experiments: (1) sitting in front of a desktop computer; and (2) walking in an indoor environment.

The contributions of this work include: (1) validation of occlusion and distraction issues in Glanceable AR, (2) design considerations for developing techniques to activate minimized glanceable information, (3) design and evaluation of five promising techniques, and (4) design recommendations for future implementations of these techniques.

This work has been published and won a Best Paper Honorable Mention Award at ACM SUI 2021.

4.2 Validating Occlusion Issues of AR Content

4.2.1 Interface Conditions

In the study, participants were asked to wear a Magic Leap One AR headset. The headset displayed four pieces of virtual information (weather, calendar, email, to-do list) in different forms that varied in LoD and location. The virtual information was not interactive—it was always displayed, so that there were opportunities for the virtual content to occlude the real world. We used two ways of locating the information in the periphery (heads-up display (HUD) or head-glance (HG)) and two LoDs (Full-app or Icon), yielding four interfaces in total (see Figure 4.1). In the HUD condition, content was display-fixed at the edges of the field of view (FoV) of the AR display and always visible. Such HUDs are commonly used in commercial products without head tracking such as Google Glass\(^1\) and North Focals\(^2\). The HG condition used a glanceable AR interface proposed by Lu et al. that showed potential for everyday uses [100]. In HG, content was fixed to the torso of users’ body and stayed outside the central view of the users. It became visible only when the user turned their head.

4.2.2 Tasks

In the study, participants were asked to walk in a 6.22 by 9.95 meter indoor environment casually and identify certain paper signs in the real world (see Figure 4.4 (b)). This task was designed to represent common scenarios in which people walk around in the real world and search visually for a target or destination (e.g., a conference room in a building with a certain room number). Twenty signs were distributed randomly in the space on walls, whiteboards, floors, and tables. Participants were asked to wear the AR headset, walk in

\(^1\)https://www.google.com/glass/start/
\(^2\)https://www.bynorth.com/
CHAPTER 4. RESOLVING OCCLUSION ISSUES OF AR CONTENT

Figure 4.1: The four interfaces in the first study with information at the periphery: (a) HG-Full; (b) HUD-Full; (c) HG-Icon; and (d) HUD-Icon.

an figure-eight path at a comfortable speed, and search for signs that contained a certain letter, which varied with every trial. Only some of the signs contained the target letter. Participants were asked to speak out loud the word in each sign that contained the letter as soon as they noticed it. Interface condition was the only independent variable for the study. We used a within-subjects design, in which each participant experienced the five interface conditions one by one. Latin square counterbalancing was applied to the order in which conditions were experienced. The task took about one minute to complete for each condition.

4.2.3 Participants & Procedures

Eight participants (7M/1F) between 19 and 33 years old (M=24.37, SD=1.64) were recruited for the study. The participants were all college students. Four participants did not have experience with AR before the study.
4.2. Validating Occlusion Issues of AR Content

The experiment was divided into five phases. First, 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, participants were asked to fill out a background questionnaire. Third, participants were given a brief introduction to the experiment background, hardware, five conditions, and the tasks involved in the study. Fourth, after participants had no further questions, they were instructed to put on the AR headset and go through the conditions one by one. After each condition, participants were asked to fill out a questionnaire containing Likert-scale questions that asked about the level of distraction and intrusiveness of the AR interface (see Figure 4.1). After finishing all conditions, a brief interview was conducted, in which we asked participants about their preferences and comments about the AR interfaces. The study took 30 minutes in total.

4.2.4 Results & Discussion

Results in the Likert ratings showed that all four interface conditions caused some visual interference (occlusion) when users were visually searching the signs (see Figure 4.2 (a-c)). The HUD-Full condition was found to be the most annoying and got in the way the most often.

In the interview, ignoring the baseline condition, six out of eight (75%) participants ranked the HG-Icon condition the most preferred and found it the least distracting. Seven out of eight (87.5%) participants disliked the HUD-Full condition, and found it the most distracting.

For the HUD-Full condition, participants commented: the icons being fully displayed greatly reduced my vision and distracted my focus, and I needed to pause to read the sign because I was trying to look past the items. Similar comments were made for the HUD-Icon condition: the icons were obtrusive, they felt like it was crowding my field of vision. For the HG-Icon
condition, participants commented: *It fits comfortably and doesn’t limit my view too much,* and *Icons are small and don’t take up much of the visibility.* For the HG-Full condition, participants commented: *I liked the interface because it wasn’t directly in my line of sight.* However, four participants (50%) felt that it was still distracting to display the full app in HG-Full. They commented: *It still gets in the way sometimes when I look for the signs,* and *I would prefer the icon one [HG-Icon] better.*

In general, our results indicate that even when content is placed in the periphery, it can still be noticeable and annoying, and it can still visually interfere by occluding the real world during visual search tasks. However, these effects can be reduced by placing the content outside the central visual field (using the HG technique) and by reducing the LoD to a less obtrusive icon. Using an icon to represent the virtual content leads directly to the need for techniques that can activate the full virtual content when the user wishes to access that
4.3 Techniques for Rapid Activation of Glanceable Content

4.3.1 Design Considerations

Having established that Glanceable AR interfaces can still lead to occlusion and distraction when examining real-world information, we next brainstormed possible methods for accessing virtual content while minimizing occlusion problems. Based on the results of the preliminary study, we decided to use a real-world prioritized strategy, where virtual content is attached to the body and represented as small, semi-transparent icons by default (similar to the HG-Icon condition) [42, 125]. We then need methods for activating the virtual content in order to display the full version, and for deactivating the virtual content when the user is finished with it. We highlight five aspects below that need to be carefully considered for designing the solution interfaces.

Default visualization: When virtual information is minimized by default, it becomes critical for the users to be aware of where the information is so that they can activate and look at the information. Results of the preliminary study shows that representing minimized virtual information by an icon could be an effective strategy to prevent the virtual information from blocking users’ view, while still maintaining some level of awareness towards where to look at to activate the content. As such, we apply the same strategy and set icons as the virtual information’s default inactive state. Upon activation, the icon would go away and be replaced by the corresponding full-version virtual content.
CHAPTER 4. RESOLVING OCCLUSION ISSUES OF AR CONTENT

Deactivation: Virtual information needs to be deactivated (minimized) automatically when users stop paying attention to it. A straightforward method is to use users’ real-time gaze direction. Gaze direction provides a good indication of where users’ current visual attention is located \[59\]. When they are no longer gazing in the direction of a piece of virtual information, we can assume that they do not need the information anymore, and thus automatically minimize the information.

Activation: Different from deactivation, detecting the intent of activation is challenging. Noticing that users are looking in the direction of a piece of virtual content is not enough to imply that they want to activate the information. It is possible that their intention is to see the real-world environment behind the content. To prevent false activation of the virtual content and occluding users’ view, we need to design suitable techniques for users to disambiguate their interest layer (i.e., real world or virtual content) explicitly and rapidly. Existing research has mainly explored explicit interactions in selection tasks \[28, 90, 123, 170\]. To our knowledge, our work is the first that looks at these interactions to disambiguate users’ interest between AR content and real-world environments, which we believe is crucial if AR information is to be integrated pervasively in daily situations.

Input modality: which input modalities have been applied for activating the virtual information. In this research, we explore controller-free interactions, which include gaze, hand, and head-based techniques to activate the glanceable content.

Depth dimension: whether depth information is considered as part of the input dimension. Occlusion is one of the strongest depth cues \[48\]. When virtual information occludes the real world, it becomes intuitive to the users that the virtual information is at a closer layer of depth while the real-world is at a farther layer of depth. Enabling users to indicate
4.3. Techniques for Rapid Activation of Glanceable Content

Figure 4.3: (a) When the user is looking in the direction of minimized virtual content, we propose five interaction techniques to activate the virtual information: (b) Fixation-Glance, in which users converge their gaze at the depth of the content; (c) Head-Depth, in which users lean backward three centimeters; (d) Hand-Overlay, in which users put their hand slightly behind the virtual content; (e) Blink, in which users blink the eye twice within one second; and (f) Dwell, in which users maintain their gaze on the virtual content for one second.

the target depth layer of interest with the techniques could be intuitive in activating the virtual content. However, indicating the target depth could also increase user workload.

4.3.2 Activation Techniques

Based on the design considerations, we propose five techniques for explicit activation of virtual information: Dwell, Bink, Fixation-Glance, Hand-Overlay, and Head-Depth. Table 4.1 shows a summary of the five techniques.

Dwell (DW)

The most basic technique to prevent false activation of virtual information is the gaze-based dwell technique [81]. It was developed to avoid the effects of eye tracking jitter and the
“Midas Touch effect.” To activate the virtual information, users gaze in its direction and maintain their gaze for one second. The content does not appear if users look away from the target before the end of the one-second period (see Figure 4.3 (f)). We consider Dwell as the baseline technique, since it is a widely adopted method to avoid false-positive activation of virtual content.

**Blink (BL)**

Blinking has been explored in the literature in object selection tasks and assistive technologies [28, 63]. Recently, Lu et al. found that blink is advantageous as compared to dwell for text-entry in VR [105]. It could be an effective strategy to explicitly activate virtual information. With the Blink technique, users activate the virtual information by looking in its direction and blinking their eyes. We need to prevent content from being activated by involuntary blinks; therefore, we use multiple consecutive blinks. However, the increased number of required blinks could induce eye fatigue and increase the activation time. We initially experimented with three blinks in a row within 1.5 seconds, which was found to induce eye fatigue with prolonged use. Through iterative testing, we found that two consecutive blinks within a one-second window produced a good balance between speed and accuracy (see Figure 4.3 (e)). A white dot is displayed below the virtual content when the first blink from the user is successfully registered. Then the user can blink again to activate the virtual information.

**Fixation-Glance (FG)**

Gaze-depth has been explored in object selections, visualizations, and information acquisition [6, 73, 100, 121, 124, 143, 155]. The Fixation-Glance technique utilizes gaze as the input
modality with the depth of gaze being considered as an extra input dimension. To activate the virtual information, users need to not only look in its direction, but also converge the gaze at the depth of the information (see Figure 4.3 (b)). As such, if users intend to look at the real-world environment behind the virtual content, virtual information does not appear. The design of Fixation-Glance closely follows the idea of natural user interfaces (NUI), in which the user operates through intuitive interactions related to everyday natural human behavior [99]. In everyday life, our gaze naturally fixates in the direction and at the depth of objects that are of interest in order to see them clearly. The Fixation-Glance interface takes advantage of this process so that mental workload is minimized. However, due to technological limitations, current binocular eye-trackers only provide accurate estimation of gaze depth when it is within two meters. We used approximately 0.4 meters as the depth to activate the AR content because real-world objects are less likely to appear at this depth, and it was easy for users to converge at this depth without discomfort. Recent research has shown that stare-of-the-art eye trackers, such as the Tobii Pro Glasses 3, could provide good gaze vergence depth estimations from 0.25 to 32 degrees with standard deviations between 1.1 degree to 2.7 degree. [46]. However, another research showed that, as the target moves further away, due to the smaller changes in users’ vergence angles [143], higher estimation error could be induced [155]. This is also supported by early research in perceptions of depth [40], which suggested that “distance or depth errors are apt to occur in distant portions of the visual field,” as noted by Gogel in 1993 [61], “because cues of depths are attenuated or are below threshold and therefore are unable to support the perception of depth between different depths at different positions.” As such, the accuracy of vergence detection could work better if the target is in close to mid-level ranges.
Hand-Overlay (HO)

Previous research on eye-hand coordination has found that eye and hand benefit from each other in selecting and manipulating objects [77, 83]. Inspired by the hand-occlusion selection technique proposed by Argelaguet & Andujar [9], the Hand-Overlay technique utilizes the hand as the input modality with the hand depth being considered as an extra input dimension. To activate the virtual information, the user looks in its direction and puts their hand slightly behind the virtual content, blocking the real-world environment to indicate their interest in the depth layer in front of the hand (i.e., the virtual information). Virtual information activates only when the gaze ray intersects with the virtual information and users’ hand simultaneously (see Figure 4.3 (d)).

Head-Depth (HD)

The Head-Depth technique utilizes the head as the input modality with the head depth as an extra input dimension. To activate the virtual content, the users looks in its direction and then leans backward to activate it (see Figure 4.3 (c)). Following the design guidelines proposed by Yu et al., we used 3 cm (~1.18 inch) as the distance threshold to activate the virtual information [170]. The idea of using head depth was first explored in adaptive interfaces such as the LoD interface [47] and the proximity-aware user interface [69]. It was proved to be faster as compared to gaze-based dwell interactions to select objects in 3D environments [170].

Trade-offs among interfaces

Table 4.1 summarizes the five activation techniques, as well as our hypothesized trade-offs related to accessibility (ease of access) and robustness (avoidance of false activation). The
Table 4.1: The five techniques to activate or deactivate virtual information and the hypothesized trade-offs.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Activation</th>
<th>Deactivation</th>
<th>Depth</th>
<th>Accessibility</th>
<th>Robustness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dwell</td>
<td>Maintain the gaze for 1 second</td>
<td>Stop gazing at the direction</td>
<td>No</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Blink</td>
<td>Gaze at the direction of the content</td>
<td>Blink twice within 1 second</td>
<td>No</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Fixation-Glance</td>
<td>+ Converge gaze at the content’s depth at the direction</td>
<td>Yes</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Hand-Overlap</td>
<td>Put hand slightly behind the content’s depth</td>
<td>Yes</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Head-Depth</td>
<td>Move the head backward in depth for 0.03m</td>
<td>Yes</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Dwell technique offers high accessibility, as the user only needs to keep looking at the content without performing other actions. However, it would be low in robustness because looking in the direction alone does not necessarily indicate that users intend to look at the virtual content. Blink improves robustness by adding two consecutive blinks as a security layer when user is looking in the virtual content’s direction. However, we blink involuntarily all the time, so false activations can still occur. Intentional blinks could also induce eye strain when used frequently in a short period of time, reducing the accessibility. Similarly, Fixation-Glance improves robustness by requiring users to converge or diverge their gaze. We are positive about its accessibility because of natural eye fixations. However, Fixation-Glance may also be more susceptible to false activation because of involuntary eye movements. Hand/head-based input modalities offer high robustness, because they require actions that are less likely to be performed unintentionally by users. However, that also lowers the accessibility and potentially makes them more intrusive to co-present others.

4.4 Evaluations in Stationary and Mobile Contexts

4.4.1 Research Goals

A primary study was conducted with two research goals: (1) to validate the trade-offs we hypothesized and (2) to explore what factors could influence the usability, user preference, and acceptance of these interfaces.
4.4.2 Participants

Fifteen participants (7M/8F) between 19 and 35 years old (M=23.73, SD=4.48) were recruited for the study. Four of them did not have prior experience with AR before the study. Participants all had corrected or uncorrected near-perfect vision.

4.4.3 Tasks

Two types of task contexts were included in the study: sitting in front of a desktop computer and walking in an indoor environment. A 5 × 2 within-subject design was used for the study with two independent variables: five interfaces (Dwell, Blink, Fixation-Glance, Hand-Overlay, and Head-Depth), and two task contexts (Sitting and Walking). For both contexts, participants were asked to wear an AR headset, and access information in either the virtual content or in the real-world environment behind the virtual content using each of the five techniques. Participants were instructed to access the information as fast as they could (see Figure 4.4).

Sitting Task Context

In the sitting task, participants were asked to sit in front of a desktop monitor placed in front of a wall surface. The wall was approximately 0.7 meters away from the participant’s head. Six blank sticky notes were attached to the wall. Text was rendered on each sticky note using the AR display (although the sticky notes represented “real-world” content, we used AR rendering so that the text could be changed for each trial) (see Figure 4.4 (a)). Around 0.3 meters in front of each sticky note, a piece of virtual content was displayed. The virtual content and the sticky note always overlapped with each other no matter where a user’s head was located. In real-world use cases, it is unlikely that users would choose to place the virtual
content and the sticky notes so that they overlap. However, we considered a sticky note to be analogous to an attention-grabbing event in the real-world environment that happens to share the same direction with the virtual information. If the virtual content is activated inadvertently by the user, it would expand and occlude users from seeing information in the sticky notes. Thus, this setup allowed us to evaluate how robust each interface condition was towards false positives.

Participants were asked to answer questions displayed on the computer screen. Participants saw four questions regarding the virtual content, and four questions regarding information in the sticky notes, yielding a total of eight questions. After participants finished answering a question, they used a mouse to click on a “Next” button to proceed to the next question on the computer screen. The AR headset and the computer were connected via network, so when users hit the “Next” button, information in the virtual content and the sticky notes in the AR headset automatically updated. As such, users always had to look at the information in order to provide the correct answer. Participants were instructed to answer the questions verbally as soon as they figured out the answer. In cases when participants failed to answer a question (e.g., when a technique did not work well for the participants), they were allowed to skip the question and proceed to the next one.

**Walking Task Context**

In the walking task, participants were asked to walk around in an indoor environment while following a virtual floating panel rendered in AR. The panel moved along a predefined path in the room with a constant speed. Based on its distance to the AR headset, the panel also changed color. Participants were asked to maintain a safe distance from the panel during the walking task (1.3-1.7 meters) so that the color of the panel stayed green (see Figure 4.4 (b)).
During walking, the AR display rendered four pieces of virtual content using the HG metaphor 0.4 meters away from the AR headset. The content was body-fixed. Each of the four pieces of virtual content was located in one direction of users’ peripheral vision (i.e., up, down, left, right), which was approximately 26 degrees horizontally and 19 degrees vertically away from the central view. The reason that we chose four instead of six as in the sitting scenario was that walking is a more cognitively heavy task than sitting. Too much information at the periphery could limit users’ awareness of the surrounding environment. By reducing the amount of virtual content, we hoped to alleviate the problem of information overload and ensure a safe walking environment for participants.

Similar to the preliminary study, in the walking task, eight physical signs were distributed in the indoor environment so that the user would encounter two signs on their left, two on their right, two above, and two below (see Figure 4.4 (b)). During walking, the AR display asked questions verbally about information in the virtual content or the signs behind the content. Eight questions were asked in total for each trial (4 directions \( \times 2 \) source (signs/virtual content)). Similar to the sitting task, when questions were asked about a physical sign in a particular direction, the virtual content would be rendered so that it overlaid the corresponding physical sign exactly. The signs were placed at depths of 1.2 - 2 meters depending on users’ speed and location during walking.

On average, each piece of activated content occupied around 10.8 degrees of visual angle horizontally and 7.4 degrees vertically, and each minimized icon occupied around 2.4 degrees horizontally and 2.2 degrees vertically on the display.
4.4. Evaluations in Stationary and Mobile Contexts

4.4.4 Apparatus

The study used a Magic Leap One AR HWD. The device has $1280 \times 960$ resolution with 50-degree diagonal FoV. The built-in head, hand, and binocular eye-tracking sensors of the Magic Leap were used to implement the activation techniques. The display of the Magic Leap is connected to an external battery/processing unit. For the walking task, participants wore a body strap to comfortably carry the battery pack with them during walking. To realize the HG interface metaphor, they also wore a waist strap with a 3D printed case to hold the Magic Leap controller to track the body torso during the walking task. The experimental software was developed via Unity 2019.3.7f1 with the SDK provided by Magic Leap.
Figure 4.4: Illustrations of the two task contexts: (a) the sitting context; (b) the walking context.
4.4.5 Measures

Quantitative: We used the System Usability Scale (SUS) questionnaire [22], Social Acceptability Questionnaire [5] and NASA TLX workload questionnaire [70] to gauge the usability and workload of each interface. We also recorded the time it took for participants to answer the questions prompted by the computer screen/AR display. We calculated the time it took for participants to start verbally saying the answer after the question was displayed on the computer screen for the sitting task, and after a question audio finished playing in the AR headset in the walking task. We also recorded the percentage of false positives (i.e., false activation of the virtual content) when the intention was to access information in the real-world environment.

Qualitative: As for qualitative data, participants were asked to comment on what they liked or disliked about using each technique. For the walking task specifically, participants were video-recorded by a Logitech C930e camera mounted high on a wall for later analysis of walking behaviors while using the interfaces.

4.4.6 Hypotheses

• H.1. In both tasks, Fixation-Glance will be the most preferred, with faster information acquisition speed and lower workload as compared to other techniques, because our eyes naturally fixate in the direction and at the depth of objects that we are interested in, which will reduce the occurrences of conscious user input and inadvertent activations.

• H.2. Dwell will be the least preferred while accessing information in the real-world, because the user will have to obtain the information within the one-second dwell time to avoid inadvertent activation of the virtual content and occlusion of the real world.
• **H.3.** The Head-Depth technique will be the least favored technique in the walking condition, because it is challenging to lean backward accurately while walking at the same time.

### 4.4.7 Procedure

The experiment, which was approved by our university ethics board, was divided into six phases. In the first phase, participants were asked to read and sign the consent form. In the second phase, 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, five 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) the fitting guide program of the Magic Leap One to determine the ideal size of the forehead-pad and nosepad; and (2) the visual calibration program of the Magic Leap One to ensure proper functioning of eye-tracking. Fifth, participants experienced each of the five conditions one by one, first in the sitting task, then in the walking task. The order of the five interfaces was counterbalanced using a Latin Square design. Before completing the experimental task in each condition, a training session was provided to get participants familiar with the interactions. After each interface condition, participants were asked to fill out the SUS and the NASA TLX workload questionnaires, and questions about what they liked or disliked about the interface. Each condition took about six minutes. Last, after finishing all five conditions in both sitting and walking scenarios, participants were asked to fill out a post-study questionnaire, in which we asked them to rank the interfaces based on their own preferences. The entire experiment took about 90 minutes. Participants were allowed to take a break anytime in between trials.
4.4.8 Results

We conducted a series of analyses to test our hypotheses and explore the trade-offs between the interfaces. We decided not to compare between the sitting and walking task because they involved different procedures and setup. As such, we separated the data based on the task, and used a one-way repeated-measures ANOVA (RM-ANOVA), with interface condition as the only independent variable for all the analyses. Shapiro-Wilk tests were applied to test the normality of the data. Friedman tests were applied when data failed the normality tests. 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 $\alpha$ level of 0.05 in all significance tests. In the results figures, pairs that are significantly different are marked with * when $p \leq 0.05$, ** when $p \leq 0.01$, and *** when $p \leq 0.001$. For simplicity, we will use the abbreviations “DW”, “BL”, “FG”, “HO”, and “HD” for “Dwell”, “Blink”, “Fixation-Glance”, “Hand-Overlay”, and “Head-Depth” interfaces for the rest of the chapter.

Task Performance

In each task, we collected data for 4 (number of questions asked) $\times$ 2 (source of content: either virtual content or real-world) $\times$ 5 (number of interface conditions) $\times$ 15 (number of participants) trials, yielding a total of 600 trials. We separated the data based on virtual-content and real-world and averaged the values for each participant in each condition, and used those averages in our analyses, leading to a total of 75 data points per source.

Figure 4.5 (a) shows the time it took for participants to answer the questions in each condition about each source in the sitting task. For accessing information in the virtual content, RM-ANOVA revealed no significant difference among interface conditions in the sitting task.
As shown in Figure 4.5 (a), for accessing information in the real-world sticky notes, a significant main effect was found on interface conditions with a large effect size \( F(1.716, 24.026) = 6.936, \ p = .006, \eta_p^2 = .331 \). Pairwise comparisons showed that DW (M = 5.96s, SD = 3.19) led to significantly longer time to access information in the sticky notes than FG (M = 3.99s, SD = 1.74) (\( p = .041 \)). No other significant differences were found.

As shown in Figure 4.5 (b), for accessing information in the virtual content in the walking task, RM-ANOVA revealed a significant difference among interface conditions with a large effect size \( F(1.997, 27.954) = 8.711, \ p = .001, \eta_p^2 = .384 \). Pairwise comparisons showed that FG (M = 2.56s, SD = 1.08) was significantly faster than HO (M = 3.89s, SD = 1.16) (\( p = .032 \)) and HD (M = 4.66s, SD = 2.38) (\( p = .011 \)). BL (M = 2.66s, SD = .57) was also significantly faster than HO (\( p = .016 \)). For the time it took to access information in the real-world signs, RM-ANOVA found no significant differences among interfaces \( F(4, 56) = 1.590, p = .190 \).

**Overall usability**

As shown in Figure 4.5 (c-d), we found significant main effects of interface conditions on the SUS score for both the sitting \((F(4, 56) = 5.040, p = .002, \eta_p^2 = .265)\) and walking task \((F(4, 56) = 9.826, p < .001, \eta_p^2 = .412)\). Pairwise comparisons show that FG (M = 52.08, SD = 22.15) received a significantly lower SUS score than DW (M = 76.25, SD = 15.16) (\( p = .017 \)) and HD (M = 80.83, SD = 14.98) (\( p = .008 \)) in the sitting task. In the walking task, HD (M = 54.39, SD = 20.38) received a significantly lower score as compared to DW (M = 75.42, SD = 13.51) (\( p = .02 \)), BL (M = 81.46, SD = 10.73) (\( p = .001 \)), and FG (M = 84.38, SD = 12.33) (\( p < .001 \)).
4.4. Evaluations in Stationary and Mobile Contexts

Figure 4.5: (a-b) The time it took for participants to answer the questions in the virtual content/real-world for the (a) sitting task, (b) walking task; (c-d) the SUS score for the five interface conditions in the (c) sitting task and (d) walking task; and (e-f) Participants’ response in 7-point Likert scale to the question “How do you like using the interface to access information in the virtual content/real-world” for the (e) sitting task and (f) walking task (±S.E.).

User preferences

Figure 4.5 (e-f) shows participants’ responses to the question “How do you like the interface for accessing information in the virtual content / real-world?” on a 5-point Likert scale. For virtual information acquisition, DW (M = 4.67) received the highest rating while FG (M = 2.93) received the lowest average rating in the sitting task. Friedman test shows significant differences among interfaces ($\chi^2(4) = 19.730, p = .001$). Wilcoxon signed-rank test showed that DW received significantly higher rating than BL (M = 3.47) ($Z = -3.025, p = .025$) and FG ($Z = -3.088, p = .020$) for accessing virtual information in the sitting task. As for the walking task, FG (M = 4.60) received the highest average rating for accessing information in the virtual content, while HD (M = 2.67) received the lowest average rating. Friedman test shows significant differences among interfaces ($\chi^2(4) = 22.673, p < .001$). HD interface
received a significantly lower rating as compared to BL (M = 4.33) (Z = -2.989, p = .028) and FG (Z = -3.190, p = .014).

For accessing information in the real-world, BL (M = 4.73) received the highest ratings for information acquisition in the sitting task, while both FG and HO (M = 4.67) received the highest rating for the walking task. DW received the lowest rating in both sitting (M = 2.27) and walking (M = 3.33). For the sitting task, Friedman test found a significant main effect of interface on ratings (χ²(4) = 27.983, p < .001). DW received a significantly lower score than BL (Z = -3.313, p = .009), HO (M = 4.47) (Z = -3.092, p = .020), and HD (M = 4.47) (Z = -3.352, p = .008). Friedman test found a significant main effect of interface for the walking task (χ²(4) = 14.373, p = .006). However, no differences were identified in pairwise comparisons.
4.4. Evaluations in Stationary and Mobile Contexts

Figure 4.6 (a-b) shows participants’ ranking of the five interfaces. For the sitting task, participants’ preferences were very scattered. There was no clear tendency towards favoring or disliking a specific interface. In contrast, for the walking task, 8 participants (53.33%) ranked FG as the most favored interface, and 8 participants (53.33%) ranked BL as the second place. Meanwhile, 10 participants (66.67%) ranked HD as the least favored interface.

NASA TLX Workload

Our analysis revealed several significant differences in the NASA TLX workload ratings. For the sitting task, FG was considered significantly worse on Mental workload \((Z = -2.779, p = .050)\) as compared to HD. For the walking task, HD was rated significantly worse than FG \((Z = -3.008, p = .026)\) on Physical workload. HD was also rated worse on the Effort category than BL \((Z = -3.119, p = .018)\) and FG \((Z = -2.817, p = .048)\).

Percentage of false activations

We counted the percentage of false-positives (i.e., a piece of virtual content was activated falsely by participants when a question was asked about information in the real-world). As shown in Figure 4.6 (c), in the sitting condition, DW \((M = 93.33\%, SD = 11.44)\) and FG \((M = 71.67\%, SD = 31.15)\) resulted in the highest false activation rate, followed by BL \((M = 30.00\%, SD = 21.55)\), HD \((M = 10.00\%, SD = 12.67)\), and HO \((M = 6.67\%, SD = 11.44)\). For the walking task (Figure 4.6 (d)), DW \((M = 73.33\%, SD = 19.97)\) resulted in the highest false activation rate, followed by FG \((M = 20.00\%, SD = 16.90)\), BL \((M = 13.33\%, SD = 18.58)\), HO \((M = 11.67\%, SD = 16.00)\), and HD \((M = 8.33\%, SD = 12.20)\).
Figure 4.7: Social acceptance rate in percentage for (a) locations and (b) audiences.

Social acceptability

Figure 4.7 (a-b) shows participants’ responses towards the social acceptability questionnaire. In general, all interfaces can be accepted by almost all participants in private situations such as home environment or when alone. BL was considered the most acceptable to use in all locations and in front of all kinds of audiences. In contrast, HD and HO were considered less acceptable to use as compared to other interfaces.

Comments on interfaces

When asked what they like or dislike about each interface, FG was praised for being fast, easy to use, and intuitive. Participants commented: *I was able to activate and close widgets very easily, and the content just appeared naturally when I look there.* However, they also
commented *more practice is needed for this technique, it would make me lose focus of my surrounding,* and *it was much harder to use in sitting than walking.* HO was praised for being *natural and intuitive,* but participants commented that *I would be worried I might bump into people,* and *it takes time to pull up my hand.* Participants liked DW because it is easy and straightforward. However, participants disliked it for always appearing and blocking the real-world objects: *the content always expands and block my vision... I had to look away to reset so that I could read the sticky notes.* Some participants also found it challenging to gaze at the fixed location while moving: *walking and dwelling on the widget simultaneously was hard.* As for BL, participants commented that *It was very easy to trigger the widgets... it requires little physical movements.* However, due to involuntary blinks, sometimes the content was triggered unintentionally: *the widgets sometimes appear without my intention, and getting the blinks to register sometimes is difficult in the sitting condition.* Participants liked HD because of its robustness: *the widgets never pop out unless I wanted them to.* However, participants disliked it for requiring too much physical movement. Participants especially considered it awkward and annoying to use in the walking task: *it required too much movement,* *I had to stop walking in order to look at the widgets, leaning forward and backwards was somewhat awkward, especially during walking.*

### 4.4.9 Summary of Findings

We summarize our findings as follows:

- FG was the most favorable interface with the fastest information acquisition speed and highest usability ratings in the walking context.

- DW has the highest false activation rates among all interfaces in both sitting and walking contexts.
• HO and HD received low false activation rates, but were considered less usable in the walking context.

• BL received high usability and social acceptance ratings in both contexts with reasonably good performance.

4.4.10 Discussion

In H.1, we hypothesized that FG would be the most favored interface with faster access speed and lower mental workload due to its natural and intuitive interaction. Our results partially supported this hypothesis by showing that most participants preferred FG over the other interfaces in the walking condition. It also resulted in lower physical workload and effort compared to HD. However, there was no clear tendency of favoring FG in sitting task. FG received higher SUS score in walking than sitting. In the sitting task, FG was considered more mentally difficult than HD. Participants’ comments showed that FG was considered more challenging to use in sitting than walking, which was surprising to us since walking requires continuous attention to the surrounding environment. The false-positive rates in Figure 4.6 (c-d) of FG between sitting and walking were also highly different. After revisiting participants’ comments and the playback, we surmise that the major reason was the different distance in the depth dimension between the virtual content and real-world. Due to limitations in eye-tracking technology, the gaze depth estimations could be inaccurate for some users. For the sitting task, the virtual content was only 0.3 meter away from the sticky notes, while for the walking task, the virtual content was 0.8-1.6 meters away from the signs. A larger distance allowed a larger safe region when the depth estimation was inaccurate or jittering.

Our results also supported H.2 by showing that HD was the least favored when accessing
virtual content during walking. HD received the highest SUS score in the sitting task, but it had the lowest SUS score and was ranked last by most participants in walking. Participants commented that they had to fully stop in order to perform the interaction, which reflected that it was hard to use during mobile situations.

Interestingly, a few participants also commented that BL was easier to use in the walking task than sitting. They mentioned *Sometimes I trigger the wrong widget when I blink in the sitting condition, which did not happen during walking.* We speculate that this was due to the difference in the number of pieces of virtual content. One issue with the BL technique was that the eye-tracking results would become inaccurate during blinks, which was also found in a recent study [105]. As such, if the safe region around virtual content was not big enough, blinks might be registered to the virtual content nearby. We had more pieces of virtual information that were much closer to each other in the sitting task. The safe region was smaller and more cluttered, so virtual content close to each other was more likely to be triggered falsely. As such, algorithms need to be applied to stabilize the gaze direction when a blink occurs to increase the scalability of the technique. Despite this problem, our results demonstrate great potential of BL. It received good SUS scores for both tasks, was ranked in the first/second place by most participants in the walking task, and received the highest ratings on almost all social acceptability categories.

Our results support our H.3 by showing that DW was the lowest-rated technique for accessing information in the real-world. It was not robust given its highest false-positive rates in both task conditions.

On the contrary, HO and HD interfaces were very robust given their low false activation rates. However, they were also less accessible in the walking condition, and were considered less socially acceptable with longer access time and greater physical workload given the amount of physical movement required. During walking, FG and BL techniques could be a
good balance between accessibility and robustness.

In general, our results did not show significant benefits of including depth as a input dimension. Techniques with depth input could still have high workload and low efficiency depending on the input modality being used.

### 4.4.11 Design Recommendations

We distill the following design recommendations based on our findings:

- While using FG, it would be optimal to have the virtual content in a close depth layer and the real-world information in a far-away depth layer.

- When the goal is to optimize robustness (for example, when attention to the real-world environment is required), HO is a good option because they induce low false activation rates. HD is also a good option, but only in stationary scenarios.

- When the goal is to optimize accessibility of virtual content, FG and BL could be suitable interfaces to consider, since they have fast access times in scenarios where false activations are unlikely.

- When the scenarios of use are likely to change dynamically in terms of activity (stationary/mobile), depth difference between virtual and real information, and location (private/public), BL would be a good option because it achieves good performance with high acceptance across multiple scenarios.
4.5 Conclusions

There are a few limitations to our work. First, although the AR HWD we used in the study has multiple focal planes, vergence-accommodation conflict may still affect the results of using Fixation-Glance [138]. Second, we did not use machine-learning models to improve gaze depth estimations. Recent research highlighted such possibilities, which could make Fixation-Glance more usable in both tasks [95, 157]. Last, the study setups for the sitting and walking condition were not fully controlled. On the other hand, we attempted to make the task scenarios ecologically valid and representative of everyday scenarios. More controlled studies could be performed in the future to explore different tasks and contexts.

In the near future, AR glasses could support everyday information acquisition by displaying information to the users through the lens. However, virtual content could occlude users from seeing objects of interest in the real-world. In this chapter, we proposed and evaluated five techniques to help address such occlusion issue through explicit activation of virtual content in the context of Glanceable AR, a real-world prioritized information display metaphor, to guarantee easy access of both virtual and the real-world information. Our results demonstrated the trade-offs of gaze, hand, and head-based techniques in sitting and walking contexts. Our results could inspire future implementations of lightweight techniques for explicit activation of virtual content in AR HWDs.
Chapter 5

Supporting Spatial Transitions with AR Interfaces

5.1 Overview

Compared to VR headsets where users are immersed in a virtual environment, AR glasses enable people to interact with their everyday physical world with digital augmentation [10]. In a typical everyday-life activity, people will need to move around to carry out different tasks, changing their information needs on-the-go. Recent research has shed light on the potential of AR glasses to support such needs in common everyday scenarios [65, 87, 100]. However, in existing state-of-the-art AR operating systems (OS) (e.g., the Magic Leap One and the HoloLens 2), AR content defaults to stay at a fixed location until users manually move or re-instantiate it. This kind of mechanism assumes that the main use cases for AR are confined in one space, limiting the mobility and accessibility of the digital content when users move around. As such, recent research has explored the possibility of carrying AR content with the users while moving. Lages & Bowman explored an adaptive walking interface in which AR windows become body-referenced and follow the users around [91]. In chapter 3 and chapter 4, we explored display-referenced and body-referenced layouts for carrying the AR content with the users [102]. The major limitation of these approaches is scalability. Because the system has no knowledge of what the users might need, it has to
5.1. Overview

bring all the AR content that the user will possibly need, which could cause information overload and distract the users. One alternative could be that we offload part of users’ input to the system. As such, the system could intelligently surface the possibly needed apps to the users. An early study by Sohn et al. found that 72% of the information needs were prompted by contextual factors such as location changes and activities to be done [144]. Therefore, location could be a universal trigger of information needs when users move between multiple spaces in the real world.

With mobile computing (e.g. smartphones, smart watches), people can access different applications and information on-the-go. In recent mobile phones interfaces, for example the iOS’s contextual suggestions, the system could suggest the applications to open or interactions to perform based on locations, the time of the day, schedules etc (see Figure 5.1). However, with AR HWDs, the system still relies very much on the users’ effort to find and open the application that is needed at that time. This poses two challenges to the users: (1) in scenarios in which users need to focus on real-world tasks with their attention and hands occupied, manual retrieval of digital apps could be cumbersome and tedious; (2) similar to how many apps we have installed on our mobile phone nowadays, users could need access to dozens of apps everyday in their AR systems. When the number of apps scales, finding and searching for an app could take more time and effort. How could we enable easier and scalable access to the digital content as users move across different environments while needing access to some information? For example, AR HWDs could also suggest potential application to be opened when users move to a different location, similar to current solutions on mobile phone, as illustrated in Figure 5.2.

In this research, we answer these questions by designing, developing, and evaluating several mechanisms to support AR users when they transition from one space to the other. To inform our design directions, we first conducted a user study to learn about the scalability issue with
Figure 5.1: Example of app suggestions on current mobile phone interfaces.

AR content. We simulated the procedures of needing to access AR content when moving between two locations with different numbers of total apps. Next, we conducted a user study with simulated real-world tasks and setups to evaluate three user interface conditions. Specifically, we are interested in exploring three interface solutions: \textit{None} – in which the users always have to manually open the app through the menu when they arrive at a new place (baseline), simulating current solutions in AR operating systems; \textit{Some} – in which the system will automatically bring some relevant apps with the users to the destination; and \textit{All} – in which the system brings all apps to the destination no matter whether they are needed or not. We evaluated the three interfaces under different contexts when various levels of workload and attention are required for the primary task in the real-world.

This work has been submitted to IEEE TVCG for review \cite{104}. A relevant foundational work, which was not presented in this dissertation, was also published, in collaboration with Dr. Yan Xu from Reality Labs Research at Meta \cite{101}. 
Figure 5.2: An illustration of a transition-aware interface, in which the system automatically brings a subset of all AR applications that are relevant to users’ task space after the user transitions to a new space to seamlessly support their task activities. Left: in the kitchen environment, the Recipe, Fitness, Notes, Clock, Weather, and Video apps were opened; Right: when the user moves to the office, the system brings the Note and Weather app, and automatically opens the Email, Calendar, Message apps to be ready to assist with productivity tasks, as well as the Reading List and Plant apps that are relevant to the office environment. Illustration by Lei Zhang - leiz@vt.edu)


5.2 Modelling User Interactions During Spatial Transition

5.2.1 User performance modeling and scalability in Interfaces

Modeling user performance is critical for developers to understand how different factors impact task completion time and make the best design choices in order to maximize efficiency. Notable user performance models in HCI include Fitts’s Law [58] and KLM/GOMS [27]. Similar models have been extended to phones [74], smartwatches [7], and in VR/AR [24, 31, 88]. In menu designs, Cockburn et al. proposed that the performance of completing an action in a menu could be separated into item pointing time and item search time, where the latter was linearly correlated with the number of candidate selectable items ($N$) in the menu interface for inexperienced users [36], which has also been demonstrated in other research that studied visual search performance [35, 76, 165]. A critical question arises: how does the interface perform with the scaling of items/applications in the interface for future AR/MR platforms? We define this as the scalability issue, which identifies how well an interface addresses the increased level of complexity and number of selectable items in user interfaces. As smart services grow in quantity and sophistication, how we design systems that handle the growing intensity of interactions between the user and their surroundings, taking into consideration the limited bandwidth and energy of the user, remains a critical issue to be addressed. In this work, we further investigate the role of scalability in future AR/MR systems, in the context of satisfying on-the-go mobile information needs of everyday users.
5.2.2 Research Questions

In the first study, we aim at exploring the scalability issues while transitioning AR content during user movements. Specifically, we are interested in the effect of the number of total AR applications on user performance of retrieving a specific app. We aim at answering two research questions in this project.

- RQ3.A. How is user performance of retrieving a target app correlated with the total number of virtual applications after transition is made?

- RQ3.B. What are the trade-offs when we bring no, some, or all apps with the users to the transition destination?

5.2.3 Interface Conditions

Consider an example scenario in the life of a typical everyday AR user. The user has a number of AR applications opened and placed in one location (e.g., the living room) (see Figure 5.3 (a)). He relies on these AR apps to quickly obtain and be aware of certain information to assist his real-life tasks. Suddenly, something comes up and he wants to head over to another location (e.g., the kitchen) for some tasks.
Figure 5.3: Example scenarios of using the three conditions: (a) the user is at the living room with a bunch of AR apps opened and placed around. He wants to head to the kitchen for some tasks; (b) None: the user arrives at the kitchen. Since it is a different environment, users loses complete access to the apps in the living room, thus having to use a menu to re-open the AR apps that he needs; (c) Some: the user arrives at the kitchen. The system opens the kitchen-related applications automatically for the users (e.g., recipe, timer, shopping list, calories). However, the user would still need to use the menu if he wants to open other apps; (d) All: the system automatically opens all AR applications and places them around in the environment, no matter needed or not.

None

The first condition, None, represents the current solution to access a piece of AR content at a new space. In the None condition, when users arrive at a new location, all virtual content is still in the old location. To retrieve a piece of AR information placed in the previous location, users have to visually search for the app in a menu and reopen it manually (see
5.2. MODELLING USER INTERACTIONS DURING SPATIAL TRANSITION

Some

The second condition, *Some*, indicates that the system is aware of the location changes initiated by the users after the transition process. As such, in the *Some* condition, the system would automatically open the applications relevant to the new location for the users. For example, if users move from a previous space to a kitchen space, the system would open the Recipe, Fitness, and Timer applications automatically and place them at the relative same positions (see Figure 5.3 (c)). Therefore, no manual interactions would be needed if users want to access any of the destination-related applications. However, it might happen that the system made the wrong choices that the app needed by the user is not automatically opened. If this happens, the users would still need to use a menu to reopen the application, similar to the *None* condition.

All

In the third condition, *All*, the system automatically brings all AR content with the users to the destination no matter needed or not. As such, users would be able to access any AR application without reopening it (see Figure 5.3 (d)). However, if the number of AR applications increases, it may cause visual clutter and increase the time for the user to find a target application.

5.2.4 Hypothesis

We hypothesize that the time to access an AR app will be significantly affected by and linearly correlated with the number of apps for all three interface conditions, because the
users will have to visually search within a larger set of apps in order to find the needed app when the number of apps scales. However, the slope (regression coefficient) and intercept would be different for each interface. Figure 5.4 illustrates our hypothesized trade-offs. If the total number of apps is $N$ and the number of suggested app in Some is $n_1$:

**Intercept when $N = n_1$:** We hypothesize that *All* will have the lowest intercept, since when the number of app is low, the user would know where to find the app without the need to open it manually. For *None*, however, the intercept would be higher than *All* given that the user will have to use the menu to open the app first. For *Some-Space*, the intercept will be the time it requires to find an app in the list of suggested apps, since no matter how $N$ scales, users would only need to search within $n_1$ apps, similar to when there were only $n_1$ apps in the *All* condition. For *Some-Menu*, however, the users would need to justify that the target app was not present in the list of $n_1$ apps before opening the menu. It adds an extra step to the *None* condition, hence leading to the highest intercept.

**Slope & Performance:** Figure 5.4 illustrates our hypotheses about the effect of the number of apps on access time in each of our interface conditions. We hypothesize that *All* will have a larger slope due to visual clutter, the limited FoV of the display, and the need to turn the head around in order to search in a larger area (see Figure 5.4 (All)). *None*, on the other hand, would have a lower slope since the menu is designed to be smaller with all the icons visible without needing to look around as much (see Figure 5.4 (None)). For the *Some* condition, it is a bit tricky as there are two possibilities. In cases where the system makes the correct suggestions (i.e., the target app is one of the list of $n_1$ apps the system automatically opens), the slope would be close to zero since users always only need to search within the $n_1$ apps for the target app no matter what the total number of app becomes (see Figure 5.4 (Some-Space)). In cases where the system made the wrong choices and the
user have to use the menu, the slope would be similar to the None condition, but with a higher intercept, as mentioned earlier. The performance of the overall Some condition would be largely dependent on the accuracy of the system prediction. The overall performance of Some (i.e., Some-Overall) would be a combination of the best-case performance (i.e., Some-Space) and worst-case performance (i.e., Some-Menu), where the weighted average depends on how many apps there are and therefore how likely it is that the target app is one of the n1 apps. When the number of app is low, the probability of the needed app being one of the n1 suggested apps would be fairly high. As such, the performance of Some-Overall with a small N would be close to Some-Space. As the number of app increases, the accuracy of predicting the needed app would downgrade. Hence, with the probability of needing to search in the menu turning higher, the performance of Some-Overall would become worse, capped by the Some-Menu line eventually.

Most efficient interface = \[
\begin{cases}
  \text{All,} & \text{if } N < n1 \\
  \text{Some,} & \text{if } n1 \leq N \leq n2 \\
  \text{None,} & \text{if } N > n3
\end{cases}
\] (5.1)

We hypothesize that the regression line of the Some-Overall will intersect with both the None and All conditions. As a result, as shown in Figure 5.4 All would be the most efficient interface when N is small (N < n1), because users could easily search within a small set of apps for the needed one. As the number of N increases, the time it takes to find an app in the All condition would scale rapidly, and Some would be more efficient given that it restricts the search list to n1 apps without requiring the users to reopen an app when the prediction was correct (n1 < N < n3). In the end, when N becomes really large, None would become the ideal interface to use, which gives users the flexibility and freedom to choose the app they need, and the menu allows the users to skim through the app list easier.
Figure 5.4: The hypothesized correlation between time it takes for the user to find the target app and the total number of app in each interface condition.

5.2.5 Tasks

AR apps. We used a Magic Leap One AR headset for the study. Participants had access to a set of AR applications on a virtual board in each of two locations (shown as squares on the floor) in the lab. The virtual board of each square (and the AR apps on the board) was only visible when the user was inside the square. To reduce the potential confounding variables, the AR apps were simulated with a red background and a random icon. They were arranged in a spiral format on a virtual board in front of the users (see Figure 5.5). We used a spiral layout in front of the user because we wanted to minimize the time users need to search and look around in space, assuming good familiarity with the layout and position of the AR apps. The total number of AR apps varied from 8 to 28 with a gap of 4 (i.e., 8, 12, 16, 20, 24, 28). We define the number of suggested apps (i.e., n1) in Some as 6, which we believe is a reasonable number for users to browse through without causing heavy mental
5.2. MODELLING USER INTERACTIONS DURING SPATIAL TRANSITION

or physical workload.

Figure 5.5: The task scenario, in which two “locations” were simulated by two squares A and B in AR. Participants were asked to move from one square to another and find a target AR application, simulating a typical transition scenario in real-world scenarios.

**Transition.** By quantifying and modelling user performance when the number of AR apps varies, we are able to understand the trade-offs among different interface solutions and study which interface would lead to the optimal performance with different numbers of apps. To simulate the spatial transition procedure (i.e., move from one space to another) in the real-world as much as we can while collecting sufficient data samples for modelling, we designed a task which requires participants to move between two “locations”, A and B, back and forth in the real world repetitively. Each “location” was simulated by a 51 inch by 51 inch square on the floor. Each square was 6 inch apart from the other (see Figure 5.5). In each trial, participants would start in either square A or B, depending on where they left off in the last trial, then transition to the other square and find the target app with the icon indicated.
on the Magic Leap One controller (see Figure 5.6 (b)). At the beginning of each trial when users were located in the first square, the virtual board shows all the AR apps to the users, giving them a sense of the total number of the AR apps and the position of each app. After the user goes to the new square, whether users could still see the AR apps on the virtual board depended on the interface condition. In the None condition, all AR apps would be invisible (see Figure 5.6 (b)), and the users had to use a menu to reopen the target app (see Figure 5.6 (c)) and click on it in the virtual board to finish the trial (see Figure 5.6 (d)). In the Some condition, the system would bring six apps with the users and show them on the virtual board at the new location (see Figure 5.6 (e)). If the target app was in the six apps, the user could click on it to finish the trial. However, if the target app was not in the six apps, the user would need to open the menu to check the list of unopened apps and open the target app manually, similar to the None condition. In the All condition, the system would bring all the apps with the users to the new location. The users would need to locate the target app and click on it with the controller to finish the trial (see Figure 5.6 (f)) .

Prior knowledge with the apps. In everyday scenarios, when mobile phone users want to open an app, they would often roughly know which screen it is and where it is located on the screen due to their familiarity with the layout of the apps. In this study, we tried to simulate such “prior knowledge of the apps” by asking the users to click through four AR apps on the virtual board in the first square, and having the target app be one of the four apps after the user move to the other square. Since the arrangement of the AR app remain unchanged in a trial, the user could locate the target app easier by recalling where it was during the clicking procedure, similar to how we would find an app on our mobile phone.

To summarize, in a single task trial, participants were asked to (1) click through four apps in the first location and try to remember what and where they were; (2) walk to the other
square; (3) find the target app indicated on the controller on the virtual board; (4) click on it with the controller to finish the trial. In the Some condition, the target app was in the six apps that the system automatically opened on the board in half of the trials (we call this condition Some-Space), while in the other half of the trials, the user had to open the app manually through the menu (we call this condition Some-Menu). The trials were randomized so participants had no idea whether the target app would be in the six apps the system suggested prior to each trial. We doubled the number of trials in the Some condition so the number of trials in each circumstance was equal to the number of trials in either the None or the All condition. As such, we collected in total 6 (number of apps) × 4 (interface conditions: None, Some-Space, Some-Menu, All) × 3 (repetitions) = 72 trials for each participant.
Figure 5.6: (a) The system highlights four apps for the users to click through before transition simulating prior knowledge with the apps; (b) the system would then ask the user to find one of the four apps after transition, an icon appears on the controller indicating the target app to be find at the new location; (c) in the None condition, the users always need to use a menu to reopen the app they need; (d) the user clicks on the opened app using the controller to finish the trial; (e) in the Some condition, the system always brings the first six apps with the users to the new location; (f) in the All condition, the system brings all apps with the users to the new location. In all the conditions, the content and layout of the apps remain unchanged before and after transition in a single trial, but is randomized between trials.

5.2.6 Participants & Procedures

We recruited thirty participants (9F/21M) from a local university with a mean age of 22 years old. Fifteen participants had little to no experience with AR prior to the study. The experiment, which was approved by our university ethics board, was divided into six phases. In the first phase, participants were asked to read and sign the consent form. In the second phase, 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
5.2. MODELLING USER INTERACTIONS DURING SPATIAL TRANSITION

detailed introduction to the experiment background, hardware, the three interface conditions, 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 the fitting guide program of the Magic Leap One to calibrate the display and determine the ideal size of the forehead-pad and nosepad. Fifth, participants experienced each of the three conditions one by one. The order of the three interfaces was counterbalanced using a full Latin Square design. Before completing the experimental task in each condition, a training session was provided to get participants familiar with the interactions. After each interface condition, participants were asked to fill out the NASA TLX workload questionnaire and a few questions about what they liked or disliked about the interfaces. Each condition took about three to six minutes. Last, after finishing all conditions, participants were asked to rank the interfaces based on their own preferences and explain their ranking choices in a short interview. The entire experiment took about 30 minutes. Participants were allowed to take a break anytime in between trials.

5.2.7 Results

We collected in total 30 (number of participants) × 6 (number of apps) × 4 (interface conditions) × 3 (repetitions) = 2160 trials. We removed the outliers that deviated by at least two standard deviations away from the average selection time in each interface condition. As such, a total of 88 trials (4.07%) were discarded, yielding a total of 2072 data points. We conducted a series of analyses to test our hypotheses and explore the trade-offs between the interfaces. For objective measures, we used a two-way repeated-measures ANOVA (RM-ANOVA), with interface condition and the number of apps as the two independent variables. A Greenhouse-Geisser correction was applied for violations of sphericity. For subjective measures, we applied Friedman tests and Wilcoxon signed-rank tests for pairwise comparisons.
We applied Bonferroni corrections for all pair-wise comparisons, and used an $\alpha$ level of 0.05 in all significance tests.

**Time.** The results from RM-ANOVA show that both interface ($F_{2,083,172.915} = 853.152, p < .001, \eta^2 = .911$) and number of apps ($F_{5,415} = 45.570, p < .001, \eta^2 = .285$) have a significant main effect on the time it took for participants to select the target app after moving to a new position. We also identified a significant interaction between interface and number of apps ($F_{9,575,794.766} = 3.373, p < .001, \eta^2 = .039$). When the number of apps increased, *None*, *All*, and *Some-Menu* led to longer search time, while the time required for *Some-Space* maintained a similar level.

![Figure 5.7](image)

Figure 5.7: The time required for participants to retrieve a target app under each interface condition with different number of apps.

Post-hoc analysis revealed that *None* yielded a significantly shorter time in general to find the target app as compared to *Some-Menu* ($p < .001$), while longer as compared to *All* ($p < .001$) and *Some-Space* ($p < .001$). *Some-Space* yielded a significantly shorter time to obtain the target app as compared to *Some-Menu* ($p < .001$) and *All* ($p = .004$). *Some-Menu* led to significantly longer time as compared to *All* ($p < .001$).
The linear regression was conducted for all interface conditions. As shown in Figure 5.8), we identified very strong linear relationships (with $R^2 > 0.9$) between the number of apps and None ($R^2 = 0.962$), Some-Menu ($R^2 = 0.937$), All ($R^2 = 0.961$) and a high linear relationship (with $R^2 > 0.6$) for Some-Space ($R^2 = 0.679$).

Figure 5.8: (a) Average time (Y-axis) it took for each condition to find the target app when different numbers of apps are present (X-axis). Strong linear relationships were indicated for None Some-Menu; and All conditions (bars indicate 95% confidence intervals). (b) The ranking of the three interface conditions by the number of participants (Y-axis).

Ease of Use & Workload. Figure 5.9 shows the average ratings for easiness and the NASA TLX workload questionnaire. Since we asked participants to only consider their experience of using the interfaces to find a target app in one single trial (to avoid the bias that Some had double the number of trials than the other conditions), the ratings were low on a 7-point Likert scale. Friedman test yielded a significant main effect of interface on the rating of easiness and mental workload. However, post-hoc pairwise comparisons led to no significant differences between the pairs.
User preference. Figure 5.8 (b) shows participants preference rankings of the three interfaces. Overall, All was selected as the most-preferred interface by 20 participants (66.67%). None was ranked as the least-favored interface by 16 participants (53.33%). For Some, 8 participants (26.67%) ranked it as the most favored interface, and 14 participants (46.67%) ranked it as the second-favored interface.

Comments. We are still in the progress of conducting in-depth coding and theme-extraction for the interview transcripts. As for initial results, below we list some of the frequent comments from the participants during the interview. For None, participants liked that “I don’t really have to memorize anything about the icons;” and “It was slightly easier to find the target app without so many apps crowding the view.” However, participants also commented that “the menu was cumbersome to use;”, “Using the laser pointer to select items in the menu was difficult at times due to their size.” For Some, participants liked that “few of the apps were already on the screen, which made it easy for me to search (for an app);” “What I liked is that sometimes the widget I wanted to use was there on the wall already;” and “this
condition felt like a combination of the best features from both previous condition (None and All).” However, participants also commented that “I disliked how this condition was more unpredictable than the other two;” and “I disliked it when I had to go to the menu to find it because it took two steps, one to scan through the smaller subset and two to open the menu and find the widget there.” For All, participants liked that “Previous knowledge of where the apps are helped greatly in finding them in the second stage;” and “the layout is consistent and static, memorizing the locations between positions A and B is quite easy.” However, participants also commented that “It felt like using a cluttered desktop; I knew where everything was, it just took a second to find everything;” “with a large number of widgets, it got very difficult and time consuming;” and “Sometimes icons would appear outside my field of view, which required more effort on my part to actually locate it in a timely manner.”

5.2.8 Discussion

Intercept & Slope. Our hypothesis about the differences in intercept were mostly correct. As shown in Figure 5.8 (a), Some-Menu had the largest intercept, followed by None, Some-Space, and All conditions. It also aligns well with our hypothesis that the Some-Menu and None conditions shared a similar slope, and Some-Space had a slope close to zero. However, it surprised us that the All had a much smaller slope than we expected. Different from what we hypothesized in Figure 5.4, All shared a similar slope with None and Some-Menu when the number of apps increased.

Trade-offs. There was no intersection between the regression lines of None and All, which means that in our study, participants would always perform better with All as compared to None no matter how the number of apps scales. Meanwhile, we also hypothesized that All and Some-Space would intersect when N equals n1 (the number of suggested apps in Some),
which was 6 in the study. However, our results show that the two lines intersected when \( N = 13.4 \). This means that when \( 6 \leq N \leq 13 \), even in the best-cases of Some when the users only needed to search within a list of 6 apps, All still resulted in a better performance. We surmise that the major reason for this was that the simulated prior knowledge lowered the time it required to find the target apps when \( N \) increased in the All condition. In the interview, 18 out of the 30 participants (60\%) commented that they felt that the procedure of learning where the target app may be located was more helpful in the All condition. Instead of having to search the entire board from center to corner, participants only had to search within four “areas” that they clicked before moving to the new location. Even though remembering the exact locations of the “areas” could become more challenging when the number of apps increases because the areas would be often located outside the FoV, knowing where the target app might be still reduced the required searching time by discarding the need to aimlessly search in random areas on the board. The fact that all apps were always in front of the users also excluded the scenarios when the users have to search around then in order to find a target app, thus leading a more ideal situation with the All interface.

The other factor we think which made the prior knowledge more helpful in All than the other two interfaces was the visual consistency of the apps prior to and after transition. One participant commented: “The best part about this condition (All) is that, since the layout is consistent and static, memorizing the locations of the apps between positions A and B is quite easy.” Such consistency in layout and visibility offered participants great help to directly apply the knowledge they learned to the new location, which would not likely be the case in actual real-world scenarios. However, in Some or None, the system only showed some apps or no app, causing visual distinctions between layouts across the two locations. As such, some additional sense-making process were needed to apply the prior knowledge learned. Our results revealed the benefits of maintaining the physical layout of the AR apps
in the space to shorten the time it required to find the target app when the number of apps increased.

The design of the menu could also be the reason why participants performed worse in the *None* and *Some-Menu* conditions. One major issue about the menu was its size. We designed the menu to be small so users could easily skim through it without having to look around too much, which, as mentioned by a few participants, posed challenges for them to pin-point an icon with the controller. The other major issue was about the arrangements of the icons in the menu. Since in this study, the apps were simulated by icons without names or semantics, the only way to find a target app in the menu would be to go through the entire list, which was less likely to happen in everyday situations given the icons would be alphabetically ordered or categorized.

**Modelling the overall performance of Some.** Given the regression result of *Some-Menu* and *Some-Space*, we were able to model the performance of *Some* overall, which is a weighted combination of *Some-Menu* and *Some-Space* depending on the probability that the needed app is one of the six apps that the system suggested. We adopted two strategies for modeling this: (1) when the apps are chosen at random, which means a larger value of $N$ would result in a lower probability; and (2) when the probability reflects the performances of state-of-the-art AI systems, such that the likelihood of the needed app being chosen is high, regardless of the value of $N$. Figure 5.10 illustrates the derived performance models for *None*, *Some*, and *All* based on the number of total AR apps.

In the first modeling approach, we assume that the system selects six apps randomly from the pool of all apps. As such, the probability of the needed app being inside the suggested list would be $\frac{6}{N}$, which would be high when $N = 8$ (75%). However, as $N$ increases, the probability would drop and became the lowest when $N = 28$ (28.57%). As such, *Some-
Overall outperforms None all the time. However, All would always be the more efficient option no matter how \( N \) scales (see Equation 5.2 and Figure 5.10 Some(Worst)).

\[
\text{Some} = \frac{6}{N} \cdot (0.009 \cdot N + 2.3) + (1 - \frac{6}{N}) \cdot (0.062 \cdot N + 5.170) \quad (5.2)
\]

In the second modeling approach, we assume the prediction accuracy to be at a fixed value. The accuracy of predictive systems proposed by previous research for predicting user needs of information or interaction intent fell around 70% [78, 130, 150, 166]. One thing different in our situation was that the system made “six attempts” by suggesting six apps in parallel. A recent study about customer service bots proposed a reinforcement learning system for suggesting candidate questions in customer service bot, in which the hit rate of the top-recommended six questions reached 95% in a candidate pool of 200 questions classes [29]. In this modeling approach, since we had a much smaller app pool, we assume that the accuracy of the needed app being one of the six recommended apps stays at 95% when \( 8 \leq N \leq 28 \). As such, the performance model of Some-Overall could be represented as in Equation 5.3. This symbolizes the best-case overall performance of Some (see Figure 5.10 Some(Best)). In this case, Some would always outperform None, and outperform All when the number of apps exceeds 17.

\[
\text{Some} = 95\% \cdot (0.009 \cdot N + 2.3) + 5\% \cdot (0.062 \cdot N + 5.170) \quad (5.3)
\]

Figure 5.10 shows the result, in which Some-Overall outperforms None all the time. However, All would always be the more efficient option no matter how \( N \) scales. As such, if the system randomly chooses the apps to transition, Some, although would be a better option all the way than None, would always perform worse than the All condition.
Overall, it appears that our result strongly favored \textit{All} as the solution for transitioning AR apps on-the-go. However, our study simulates an ideal situation of the \textit{All} condition, in which we assumed (1) all AR apps were directly in front of the user without the need to search around; (2) the layout of the AR apps remained exactly the same before and after the transition; (3) the user was not required to pay attention to the real-world environment that could be occluded by the AR apps. The \textit{Some} condition, although performed worse than we thought, still has merit if the number of apps is larger than a certain number when the system has perfect or nearly-perfect accuracy. Therefore, accuracy plays a huge role in the usability of the \textit{Some} condition, as a higher accuracy reduced the penalty of having to bring up the menu to retrieve the app manually. One reason of the penalty being high in our study could be the design of the menu, as reflected by most participants. The menu has to be easy-to-use so the users could easily recover from the system errors when it happens. Finally, \textit{None}, as a representative of existing solutions, always performed worse than the other two interfaces. Our study successfully demonstrated that the trade-off exists among
the three solutions to transition AR apps. With the differences of slope and intercept of each condition, each interface holds characteristics that could make it more usable in a certain situation. However, a more ecologically-valid study will be required to further reveal such trade-offs.

5.2.9 Conclusions

Overall, our results successfully demonstrated that the trade-off exists among the three interfaces when the number of applications differs. It appears that our result strongly favored All as the solution for transitioning AR apps on-the-go. However, our study simulates an ideal situation of the All condition, in which we assumed (1) all AR apps were directly in front of the user without the need to search around; (2) the layout of the AR apps remained exactly the same before and after the transition; (3) the user was not required to pay attention to the real-world environment that could be occluded by the AR apps. The Some condition, although performing worse than we thought, still has merit if the number of apps is larger than a certain number when the system has perfect or nearly-perfect accuracy in predicting the needed app. Therefore, accuracy plays an important role in the usability of the Some condition, as a higher accuracy reduced the penalty of having to bring up the menu to retrieve the app manually. One reason for the penalty being high in our study is the design of the menu, as reflected by most participants. The constant needs of calling the menu through a button press and using Raycasting to select small targets made the interaction tedious and time-consuming to perform repetitively. The menu has to be easy-to-use so the users can easily recover when the system does not predict the needed app. Finally, None, as a representative of existing solutions, always performed worse than the other two interfaces.

Our study successfully demonstrated that the trade-off exists among the three solutions to transition AR apps. With the differences of slope and intercept of each condition, each
interface holds characteristics that could make it more usable in a certain situation. However, a more ecologically valid study is required to further reveal such trade-offs, especially in the situation when the number of needed apps is reasonably large.

5.3 Evaluating Realistic Transitions In-Situ

5.3.1 Research Questions

In the first study, we revealed the performance trade-offs of each condition and the importance of the menu design for optimized user experience in the None and Some conditions. In the second study, we aim at exploring the three interface solutions in more ecologically valid scenarios when the number of AR applications is reasonably large. We aim at answering the following research questions:

- **RQ3.A.** With a relatively large number of virtual applications, how well does each interface support information access tasks?

- **RQ3.B.** How do different contextual factors affect the user experience with and preference for the three interfaces?

- **RQ3.C.** What are the trade-offs among the three interface conditions in ecologically valid simulations of everyday scenarios?
5.3.2 Interfaces

**Improvements to the interfaces**

Based on the lessons learned from the first study, we made a few improvements to the menu interface. Instead of using the controller to select the small icons in a world-referenced menu, which was cumbersome and time-consuming, we adopted eye-tracking for selection and a hand-referenced menu. Users could simply gaze at an icon in the menu to interact with it. The menu appears next to the hand when the hand is in view (see Figure 5.11 (b-d)). To avoid the “Midas touch effect [82]”, we adopted a 0.5 second dwell time. After gazing at an icon in the menu for more than 0.5 second, the application would be spawned from the menu location and smoothly move to its assigned location in the space. As such, the user could perform quick interactions with the AR apps without needing both hands, which would be convenient in situations when users need to have their hands occupied by tasks in the real world. In addition, we displayed the apps in alphabetical order in the menu to make them easier to find.

We used the same three interface conditions as in study 1:

- **None**: After a user transition from one location to another, users would lose access to all the apps they opened in the previous location. Apps had to be opened manually from the menu, and the icons of opened apps became opaque in the menu.

- **Some**: After a user transition from one location to another, the system would automatically open the apps that are relevant to the new location. The new apps opened by the system would have their icons become opaque with a blue star on one side to indicate
5.3. Evaluating Realistic Transitions In-Situ

Figure 5.11: (a-b) The user raises the hand to summon the menu, the currently opened/closed apps were shown as opaque/transparent icons, and the system automated apps have a blue asterisk mark on the side; (c) the user gazes at the icon for 0.5 second to open an app manually a white boarder graduate appears around the icon as visual feedback for the dwell interaction (the icons in the menu are alphabetically ordered for easy search); (d) the icon turns opaque, indicating that the app has been opened in space; (e) For Some, the icons of the opened apps were shown on the top of the field of view.

that they had been opened automatically by the system (see Figure 5.11 (b-d)). The users could make edits to the system-suggested apps by closing them or opening new ones. To help the users keep track of what apps are already opened and what are not, as commented by participants in study 1, we added visual indications at the top of the FoV. The apps opened in the current space will have their icons displayed there for quick glances (see Figure 5.11 (a-e)).

- **All**: After the user arrives at a new location, the system would automatically open and place all applications. The user could make edits by closing the apps with the menu. However, all apps would still appear automatically at the next location.
5.3.3 System Implementation

The prototype was implemented on the Microsoft HoloLens 2. A total of twenty AR apps were integrated into the system, including *Appliance*, *Book List*, *Calculator*, *Calendar*, *Chrome*, *Clock*, *Email*, *Facebook*, *Fitness*, *Map*, *Message*, *Music*, *News*, *Notes*, *Plant*, *Recipe*, *Sports*, *Stock*, *To-do List*, and *Weather*. Each app consisted of an icon, some images and/or a few lines of text. Each app had its own predefined position at each location, which was designed to be consistent across the three spaces. For example, the calendar app would always be located on the right side after the user arrived at a new location facing the entrance orientation. The reasons for this design choice were two-fold: (1) an app would be placed automatically after being opened without the need for the users to manually place it, which has been proven to be important while relying on AR UIs for real-world tasks [101]; (2) the knowledge that users gain about where the app was located in one location applied to all the other locations to maximize learnability and minimize the required mental effort, as shown in the first study. Users were not allowed to change the location of the AR apps, as our study focused on the spatial transition rather than the placement or arrangement of AR content.

5.3.4 A Contextual Framework and Hypothesis

To study the user experience of these interfaces in a deeper and more nuanced way than in study 1, we designed study 2 to examine the influence of three major contextual factors that we hypothesized would impact the user experience:

- **F1: Real world (RW) Priority.** The first factor identifies whether occlusion, distraction, or cluttering issues of the AR apps is a problem in the current context, which is defined by how prioritized the real world is in a given context. If the real world
5.3. Evaluating Realistic Transitions In-Situ

Figure 5.12: The design of the menu to make users clearly aware of (1) which app is opened/closed in the current location; (2) which app is opened by the system automatically for the current location (only in the Some condition indicated by a blue star); and (3) which app is pinned by the user and will be automatically opened in the next locations (indicated by a pin icon). The apps are arranged alphabetically to reduce the visual search effort.

is prioritized, which means that the task requires careful monitoring or engagement with objects in the real-world environment, users may only want to open the necessary AR applications to avoid occlusion, distraction, and visual clutter [42, 125]. Hence, the None may be more favored than Some and All. If the real world is not prioritized, which means that the current task does not require users’ attention on the real-world environment, then users would be less concerned if more AR apps are present in the physical world. As such, the Some or All conditions may be ideal.

- **F2: Workload.** The second factor relates to how heavy the workload is for the users’ current task, indicating the bandwidth users have available to interact with the AR applications. In lightweight scenarios when the primary task is simple, the users would have high cognitive bandwidth available, thus they may not care even if they need to spend more effort to open an app manually. As such, all three interfaces may work
well. However, if the user is engaged in heavy tasks in the real-world or the task is time-sensitive, users may have little cognitive bandwidth available. In such situations, they would want to minimize the effort needed to perform manual interactions to open an app, which brings advantages to the Some and All conditions.

- **F3: Specificity of needs.** The third factor refers to the availability and relevance of the AR apps to the tasks at hand, which also means how specific the purpose of the task is. If the user has a primary task with a specific purpose, which means that users have a good idea of what apps they need for the given task at the moment, the availability and relevance of the AR apps would be at stake to the users because the tasks highly depend on the apps. In such cases, the Some interface may be best, since it provides access to the most relevant apps for the current location automatically. However, if the purpose is non-specific (e.g., when the users has no specific task (i.e., just ”hanging out”) but may want to browse some AR information if they think of it, or where the user has a task that is unrelated to any AR app, but may want to browse some AR information as a secondary task if they think of it), then the None and All conditions may be best, since they allow the user to open or browse any app as desired.

Table 5.1 shows the hypothesized performance of each interface condition under each contextual factor. For F1, occlusion/distraction/clutter would become worse from None to Some to All, as more AR apps would become present by default after the user arrives at a new location. Regarding F2, None condition becomes the worst option because the users always need to manually open an app (high physical workload). All would be slightly better than None since the user no longer need to open any apps (low physical workload), but still we hypothesize that it will be challenging for the users to pinpoint a specific app when all apps are present in the physical world (high mental workload). Instead, Some would be a middle-ground between the two given that the number of opened apps is limited and the
users only need to open an app when it is not relevant to the current location, resulting in low mental workload and frequently-low physical workload. Finally, regarding F3, None will still be the worst given that no app is available by default. All makes all apps available while at the cost of bringing irrelevant and unnecessary apps, resulting in high availability but low relevance. Some would instead have a medium availability because a subset of the apps are opened. However, the apps that are automatically opened will be highly relevant to the current location.

Based on the three factors, we defined a framework for categorizing users’ context after transitioning to a new location, leading to eight contextual situations, to help us evaluate the impact of these factors on the interface preference in a more balanced and ecologically-valid form.

**C1: RW Not Prioritized + Light + Non-Specific:** The user randomly browses non-specific AR apps without needing to pay attention to the RW environment (for example, casually browsing some random AR apps to kill time).

**C2: RW Not Prioritized + Light + Specific:** The user needs access to some specific apps without needing to pay attention to the RW environment (for example, the user wants to know the weather for today or the number of steps walked in a relaxed setting).

**C3: RW Prioritized + Light + Non-Specific:** The user focuses on a simple task in the RW

---

**Table 5.1: Hypothesized contextual preferences on interfaces.**

<table>
<thead>
<tr>
<th>Factors</th>
<th>None</th>
<th>Some</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>RW-Prioritized</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RW Not Prioritized</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>High Workload</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Low Workload</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Specific Needs</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-specific Needs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
and sometimes wants to know non-specific or irrelevant information from AR apps. Examples include casually walking or waiting for food in the microwave while browsing some random AR apps.

**C4:** *RW Prioritized + Light + Specific:* The user focuses on a simple task in the RW while needing access to specific and relevant AR apps to assist the task. Examples include grabbing the ingredients in the fridge according to a virtual recipe or walking on the street according to a virtual map.

**C5:** *RW Not Prioritized + Heavy + Non-Specific:* The user engages with a heavy digital task while wanting to access non-specific or irrelevant AR applications at the same time. An example would be playing video games with a virtual monitor while monitoring the scoreboard of a football match in the sports app.

**C6:** *RW Not Prioritized + Heavy + Specific:* The user engages with a heavy digital task while wanting to know specific and relevant AR applications at the same time. Examples include performing productivity work in front of a virtual monitor while needing information from the notes or the browser app to assist the work.

**C7:** *RW Prioritized + Heavy + Non-Specific:* The user engages in a heavy task which requires continuous attention to the RW while needing access to non-specific or irrelevant AR content. Examples include cleaning the house while browsing the news or monitoring the stock prices in AR.

**C8:** *RW Prioritized + Heavy + Specific:* The user engages in a heavy task which requires continuous attention to the RW while needing access to specific and relevant AR content to assist the task. Examples include cooking according to a virtual recipe, or assembling furniture according to digital instructions.

We excluded *C5* and *C6* from our exploration, given that they focus on primarily the
interactions with digital tasks, for example doing productivity work with virtual monitors or playing immersive games, with neither the need nor the cognitive bandwidth available for real-world interactions, which happens less frequently during transitional activities. As such, in this work, we focus on $C_1$, $C_2$, $C_3$, $C_4$, $C_7$, and $C_8$ to explore how the three contextual factors $F_1$, $F_2$, and $F_3$ could affect user experience and preference of using the three interface conditions under these contextual scenarios.

5.3.5 Scenarios & Tasks

To understand how these contexts impact the user experience of the three interfaces in AR, we designed a user study with three goals: (1) the task needs to simulate everyday scenarios that involve on-the-go information needs with high ecological validity; (2) the task needs to happen in a realistic AR setup in a scenario that the users are familiar with; (3) the task needs to cover the six contexts in a balanced manner.

As such, we designed a study which involves embodied transitions in a large multi-room physical space while wearing an AR HWD. Participants were asked to go through a set of role-playing scenarios in the real-world environment while using the three interface conditions to obtain information. In the scenarios, participants played the role of an everyday AR user who moves around their own “home environment” for everyday tasks while needing discretionary access to some AR apps on-the-go. The “home environment” was a public indoor area located on our local campus. It consisted of a kitchen, a living room, and a home office (see Table 5.2 and Figure 5.13). In each scenario, the user will be asked to start in one of the three locations, then transition three times, each to a different location, and access some AR apps at the new location. To cover the three contextual factors we mentioned earlier, we adopted different combinations of the six contextual situations in each
CHAPTER 5. SUPPORTING SPATIAL TRANSITIONS WITH AR INTERFACES

Table 5.2: Characteristics of the three physical rooms used in the study and the list of automated applications in the Some interface.

<table>
<thead>
<tr>
<th>Room Space</th>
<th>Physical objects in Room</th>
<th>Automated AR applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office</td>
<td>laptop, desk, chair</td>
<td>Email, Calendar, Message, Clock, Weather</td>
</tr>
<tr>
<td>Kitchen</td>
<td>fridge, coffee maker, sink, kitchen counter</td>
<td>Recipe, Fitness, Note, Clock, Weather</td>
</tr>
<tr>
<td>Living Room</td>
<td>TV, plant, bookshelf, table, chair</td>
<td>News, Stock, Reading List, Plant, Clock, Weather</td>
</tr>
</tbody>
</table>

Figure 5.13: Deployments of the AR apps (i.e., the All interface) in the three physical room spaces for the transition task in study 2: (a) office, (b) kitchen, and (c) living room.

scenario (i.e., C1-C4, C7, C8). Each scenario included three out of the six contextual situations, each triggered after a transition. As such, a set of four scenarios would lead to twelve contexts, covering each of the contextual situations two times in a balanced form.

To motivate the participants to transition to a different location, refer to information in the AR apps, and conduct real-life tasks without breaking immersion of the in-situ role-playing experience, we integrated voiceovers of both third-person and first-person narrations as instructions to guide participants on what room to go to next and what task activity to perform. For example, for first-person narrations, participants would hear “It has been a long day of work! I’d love to make a coffee in the kitchen;” or “Oh no, the guest will arrive within an hour. I need to clean the rooms as fast as possible!” For third-person narrations, which usually followed right after a first-person narration, participants would hear “You head over to the kitchen and make a cup of coffee;” or “You search for trash in the room and collect it.” Please refer to the supplementary materials for the four scenarios developed for the study. The order of testing the interface was balanced through a Latin-Square. To mitigate potential learning effects, the order of experience the four scenarios was randomized for each
interface. We also randomized the displayed content in the AR apps at the beginning of each new scenario, and asked participants to follow the narration and check the apps each time they were asked to. Below shows an example scenario:

**Starting - Office:** You are working in your home office, with *email, calendar, message* app opened around your desk.

“Ah it’s been a long day of work! It’s time to make a coffee in the kitchen. I’ll also have to pay attention to my *email* and *message* app, just in case my colleague sends me the documents I need.”

**T1: Transition 1 - Kitchen:** You go to the kitchen and starts making a coffee while monitoring your *email* and *message* apps. Suddenly, you saw the fridge in the kitchen environment

“Oh right Let’s see what I need to pick up at the grocery store later for the recipe I bookmarked earlier in the *recipe* app”

You open the fridge with the *recipe* app and shopping list in your *note* app, and cross-check what ingredients you’ll need to pick up later in the grocery store.

“Done! Let’s go to the living room and take some rest.”

[* C4: RW Prioritized + Light + Specific]*

**T2: Transition 2 - Living room** You head over to the living room and sit down at the couch.

“Now let’s browse some random stuff to kill time. I wonder what is the stock price of Apple now.. oh and I have not read the trending news for today...”

You start reading the *stock* and the *news* app.

“Oh it seems that I get the documents I need! Now time to get back to work.”
CHAPTER 5. SUPPORTING SPATIAL TRANSITIONS WITH AR INTERFACES

[☉ C1: RW Not Prioritized + Light + Non-Specific]

T3: **Transition 3 - Home Office:** You head to the office and work on the documents. After it is finished, you want to head over to the grocery store to grab the ingredients you marked earlier. You look outside the window and it is a bit cloudy.

“It seems that it is going to rain soon. Should I go to grocery store later? What is the chance of rain?”

You look at the **weather** app to check the chance of being rainy.

[☉ C2: RW Not Prioritized + Light] + Specific

End

As such, this scenario touches **C1, C4, and C2**, in which users were asked to access information in the AR apps within different **contextual situations** after they move to a new location. Currently, we have developed four scenarios for the participants to experience through. We are also working on including more scenarios to encompass a broader set combinations to reveal the potential trade-offs among the three interface conditions.

**Prediction Accuracy in Some**

**Table 4.1** shows the automatically opened applications for the **Some** interface condition in each location in our prototype. The Clock and Weather applications were automatically opened in each location, due to their versatile uses and frequent needs in everyday cases. Other than that, each space had 3-4 applications relevant to the location and the tasks that happened there, and which would automatically open when the user entered that space with the **Some** interface.
In the four scenarios we designed, we simulated good prediction accuracy for the Some interface, backed up by recent successful research in predicting information needs of users. Among the twelve transitions with information queries in our four scenarios, three of them (25%) Some did not provide the application required; thus, the users had to manually open the app with the menu, simulating a 75% accuracy level. We wanted to explore how users perceived their experience with the Some interface with relatively good accuracy.

**Participants & Study Procedure**

We recruited twenty-seven participants from a local university. P1 was used as an initial pilot, and P11 & P27 were discarded due to incomplete data logs caused by software issues. The study ended up with 24 participants (4F/20M) with a mean age of 21 years old (SD = 1.13) with complete data. Fourteen participants had little to no experience with AR prior to the study.

The experiment, which was approved by our university ethics board, was divided into six phases. In the first phase, participants were asked to read and sign the consent form. In the second phase, 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, the concept of AR, the hardware we used, the three interface conditions, 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 the eye-tracking calibration. Fifth, participants were asked to open the experiment program and were provided a detailed training session to get familiar with the AR applications, their spatial locations, and how to open/close an app using the menu interface. Participants were guided to have an initial with each of the the three interfaces in each room space. They were encouraged to spend extra time in the Some
condition, in order to learn the locations of the applications that the system would automate. The reasons for this were two-fold. First, it would be challenging for participants to remember the locations of all twenty applications from the training session. For ecological validity, we wanted participants to hold partial knowledge of the apps’ locations, similar to people’s knowledge of app locations on a mobile phone. Second, the apps in Some have a higher probability of being needed, given that location is often the trigger of information. In the training, participants were told that “Here are some applications that you will likely need later, so it would be helpful if you have some idea where they are located,” to avoid biasing them towards a specific interface condition.

After finishing the four scenarios for each condition, participants were asked to fill out a post-study questionnaire with a few questions about what they liked or disliked about the interfaces. Each condition took about 10 to 15 minutes. Last, after finishing all conditions, participants were asked to rank the interfaces based on their own preferences and explain their ranking choices in a semi-structured interview. The entire experiment took about 90 minutes. Participants were allowed to take a break anytime after they finished a scenario, and were encouraged to do so after they finished each interface condition.

Measures

Objective measures. To evaluate user performance on the tasks, we computed: (1) the task completion time, i.e., how long it took a participant to finish all four scenarios using an interface condition; (2) distance travelled: how far the participant walked on average in the study sessions of each interface; (3) amount of head rotation: how much the user rotated their head each second, which could reflect how much effort users spent looking for content in AR and/or the physical surroundings.
5.3. Evaluating Realistic Transitions In-Situ

*Quantitative subjective measures.* We used the System Usability Scale [22] and two questions from the NASA TLX workload questionnaire [70] to gauge the usability and workload of each interface. We also asked participants to rate their level of agency with each interface using three questions adapted from Tapa et al. [149] to gauge the users’ sense of being in control over the AR system. We also measured how accurate they perceived the system automation to be for the *Some* and *All* conditions using a single question “I felt that the automated applications were accurate when I arrived at a new room.”

*Qualitative subjective comments.* In the interview, participants were asked to comment on what they liked and disliked about each interface condition. Additionally, they were asked to dive deeper and think about how each of the contextual factors *F1: RW Priority, F2: Workload, and F3: Specificity of Needs* impacted their preference.

5.3.6 Results

Similar to the first study, we used a one-way RM-ANOVA, with interface condition as the independent variable. A Greenhouse-Geisser correction was applied for violations of sphericity. For subjective measures, we applied Friedman tests and Wilcoxon signed-rank tests for pairwise comparisons. We used an $\alpha$ level of 0.05 in all significance tests, with Bonferroni corrections applied to all pairwise comparisons.

Task completion time & distance travelled

*Figure 5.14* (a) shows the average time it took for a participant to finish completing the four scenarios in an interface condition. *None* took the longest overall ($M = 865.67$, $SD = 551.42$), followed by *Some* ($M = 747.28$, $SD = 228.80$) and *All* ($M = 693.28$, $SD = 135.50$). RM-ANOVA yielded no significant main effect of interface on the task completion time.
CHAPTER 5. SUPPORTING SPATIAL TRANSITIONS WITH AR INTERFACES

Figure 5.14: (a) The average time it took for participants to complete all four scenarios for an interface condition (in seconds); (b) the average head rotation (in degrees/second); and (c) the average System Usability Scale (SUS) score (±S.E.).

\[ F_{1,202,27,645} = 2.546, p = .117. \]

Regarding distance travelled, participants travelled 235.08 units on average for All (SD = 4.31), followed by Some (M = 231.12, SD = 5.44) and None (M = 230.428, SD = 3.63). RM-ANOVA found no significant main effect of interface on the distance travelled \( (F_{2,46} = .569, p = .570) \).

**Head rotation**

We computed the average rate of head rotation for each of the interface conditions (see Figure 5.14 (b)). RM-ANOVA yielded a significant main effect of interface on the mean head rotation angle with a large effect size \( (F_{2,46} = 18.276, p < .001, \eta^2_p = .443) \). Post-hoc pairwise comparison showed that All (M = 4.80, SD = .21) resulted in significantly more head rotation than Some (M = 4.33, SD = .18) \( (p < .001) \) and None (M = 4.06, SD = .21) \( (p < .001) \).
5.3. Evaluating Realistic Transitions In-Situ

Figure 5.15: (a) The average ratings for the three questions regarding Agency; (b) the perceived accuracy level of the automated apps (±S.E.); and (c) the overall preference rankings of the three interfaces.

Usability & workload

Figure 5.14 (c) shows the SUS scores for each interface. Some received the highest usability score (M = 84.79, SD = 1.74), followed by None (M = 80.21, SD = 2.68), and All (M = 70.83, SD = 2.40). RM-ANOVA identified a significant main effect of interface on the SUS score with a large effect size ($F_{1,604.36.889} = 10.343, p < .001, \eta^2_p = .310$). Post-hoc pairwise comparisons showed that Some received a significantly higher usability score as compared to All ($p < .001$). The difference between None and All was borderline significant ($p = .066$).

Regarding workload, our results indicate that participants considered all three interfaces on the lower end for both physical and mental workload (all average ratings were less than 2.5). The Friedman test yielded a borderline significant main effect of interface on the physical workload level ($\chi^2(2) = 7.098, p = .076$). Wilcoxon signed-rank tests indicated that Some was considered less physically demanding than None ($Z = -2.686, p = .021$). No difference was found between None and All ($Z = -1.039, p = .897$) or Some and All ($Z = -1.931, p = .159$). The Friedman test yielded no significant main effect in terms of mental workload ($\chi^2(2) = 3.528, p = .171$).
Perceived agency & accuracy

Figure 5.15 (a) shows the average agency rating for each interface. The Friedman test revealed a significant main effect of interface on agency scores ($\chi^2(2) = 7.708, p = .021$). Wilcoxon signed-rank tests found that the None condition led to a significantly higher level of agency than All ($Z = -2.620, p = .027$). No difference was identified between None and Some ($Z = -2.201, p = .084$) or Some and All ($Z = -1.303, p = .576$).

Figure 5.15 (b) shows the perceived accuracy level of the system automation in Some and All. A Wilcoxon signed rank test found that participants considered the automated applications after transitioning to a new space to be more accurate in Some than All ($Z = -2.049, p = .040$).

Overall user preference

Figure 5.15 (c) shows the preference ranking distribution of the three interfaces among the 24 participants. Our results show that the Some interface was favored by most participants (66.67%), with only one vote as the least favorite. In contrast, 19 out of 24 (79.17%) participants ranked All as the least favored interface, with only one vote as the most favorite. 13 participants (54.16%) ranked None as the second favorable interface.

When participants were asked what they liked about None, frequently appearing themes included easy/intuitive to use (8), being in control (7), and less cluttering or occluding the real world (5). However, participants disliked having to raise hand/reopen applications each time (9) and the wait in the dwell (3).

For the Some interface, participants liked automatic opening of relevant/necessary apps (12), semantic connections (9), ease of use/helpfulness (7), and not cluttered (3). However, it was
Table 5.3: The number of votes each interface received under each contextual situation, with the interface condition with the most number of votes bolded.

<table>
<thead>
<tr>
<th>Contextual Factors</th>
<th>None</th>
<th>Some</th>
<th>All</th>
<th>No Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>RW Prioritized</td>
<td>17</td>
<td>13</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RW Not Prioritized</td>
<td>3</td>
<td>16</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>High Workload</td>
<td>11</td>
<td>12</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Low Workload</td>
<td>6</td>
<td>12</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Specific Needs</td>
<td>4</td>
<td>21</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Non-specific Needs</td>
<td>13</td>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total Votes</strong></td>
<td><strong>54</strong></td>
<td><strong>82</strong></td>
<td><strong>14</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>

criticized for sometimes blocking the real world (4).

For All, participants liked it because apps were already accessible (5). However, it received criticisms on being cluttered/overwhelming/distracting (14), covering the real-world environment (12), and opening apps that are irrelevant to tasks/space (5).

**Contextual preferences**

Other than the general preferences, in the semi-structured interview, participants were asked to dive deeper and think about how each of the contextual factors (F1, F2, and F3) impacted their preference. We organized participants’ votes on the three interfaces when the three contextual factors change (see Table 5.3).

**F1: RW Priority.** When asked whether having to pay attention to the real-world or not in the task had an impact on the interface preference, 23 out of 24 (95.83%) participants gave a positive response.

In the RW-Prioritized situations, which in the scenarios required searching for or collecting objects placed in the physical world, None received 17 votes (70.83%), followed by 13 votes (54.16%) for Some. The All condition did not receive any votes. Participants favored None because it offered clear real-world visibility. Comments included: “The fact that they were out
CHAPTER 5. SUPPORTING SPATIAL TRANSITIONS WITH AR INTERFACES

of view unless I opened them, made it so much easier to find objects or pick up trashes (P2);” and “If I have to do tasks in the real world, I would prefer to have minimal applications open (P23).” The common themes for not choosing All included visibility, occlusion, and clutter. For example, participants commented: “For All it was annoying because I need to look for stuff in the RW but it was all taken up by the AR apps (P15);” For Some, participants who voted for it did not find it too intrusive to the task and think it is acceptable to use: “I did not find it that obstructive (P7);” “I feel like the Some was good ... it did not block too much of my vision and I could still keep track of the other stuff [information in the apps] I may want to know at the circumstance (P22).”

In the RW-Not-Prioritized situation, which in the scenarios were tasks that did not require monitoring/visual searching in the physical world, Some received the most votes (66.67%), followed by All (25%), and None (12.5%). The majority of participants felt that it was beneficial having at least some of apps automatically opened if less attention is needed on the physical world. They commented: “When I do not need to interact with the real world too much, I would prefer the Some condition more, because it was very convenient (P3);” “I like, for example, in the living room having weather and clock open because it is something I would always use (P20).” Participants liked All because they thought it was easy, always available: “If I wasn’t gonna have to interact with the RW a lot, I would choose the All one, because it is easy and everything is there already (P2);”. However, participants who did not vote for All did not share this opinion: “I think I always need to pay some attention to the real world ... if there is one day that I do not want to leave my room and only get food in the fridge, I would maybe have, not All, but a lot of the apps on that day (P13);” and “I would still prefer Some, because All is too much unnecessary apps (P25).”

F2: Workload. When asked whether current workload of the task (high vs. low) had an impact on the interface preference, 20 participants gave a positive response (83.33%).
The other four participants did not consider workload to have an impact on their preferred interface.

In high-workload situations, which in the scenarios required participants to walk around, look on the bookshelf, and grab objects within a time limit, Some received the most votes (50%), followed by None (45.83%), and All (8.33%). Participants who voted Some favored its system automation: “In heavy duty task, I wouldn’t have the mind to open the app by myself (P4);” “I am running around, it does save that time to have Some. I would prefer to have the app already there (P3);” “When I have a lot more going on ... spending the time on opening the apps would not make sense (P16).” Participants who voted for None felt that it was less distracting and does not get in the way in demanding situations: “In more heavy scenarios ... I would not want the distractions from the apps to my activity (P18);” “If there is window in the way when I am in a hurry, I would get really frustrated ... In real life, without this technology ... if I want to check my email, I would go to my laptop and do that [open apps manually]. So I think I am already getting used to that, it is a traditional experience (P24).”

In low-workload situations, where participants were relaxing or conducting simple tasks, Some received the most votes (50%), followed by None (25%), and All (13.33%). Most participants still did not feel like using All, because of its distractions and clutter. Participants who liked Some favored its convenience and suggestions: “for lightweight task, I would still use Some for convenience (P13);” “When I am in a relaxed way with no activity, I would prefer the Some condition because it has a lot of options around (P20);” “I could see myself using the All condition if I was just relaxing ... but still too much to look at. I would still stick with Some, check those suggested apps and pull up the stuff I need (P23).” Participants who voted for None reasoned that they think they would have the bandwidth to open the app manually in non-demanding cases: “in lightweight case, I would be okay using None
because I can easily open the app when I need (P19)."

**F3: Specificity of needs.** When asked whether the specificity of the task had an impact on interface preference, 23 participants gave a positive response (95.83%). One participant did not consider their preference to be influenced by this factor.

When the purpose was specific, where participants needed specific apps such as needing the recipe app to pick out ingredients from the fridge, 21 out of the 24 participants (87.5%) voted for the *Some* interface, 4 (16.67%) voted for *None*, and no participant chose the *All* interface. The keywords for choosing *Some* over *None/All* included *relevant* and *natural*. Participants commented: “In cases that the apps are specific and relevant, like check the fridge with the recipe app, the apps are super useful. I like having them always there (P2);” “I would prefer having the relevant items be brought up instead of having all of them open in the specific use case (P23);” and “Some felt like the Siri suggestion on my phone, it suggested apps I am likely gonna need for my task, which is pretty cool when it works well (P24).” In contrast, *All* was criticized for being *distracting* and *irrelevant*: “Having non-relevant apps open distracted me (P20);” and “I see the calculator [app] in the kitchen [for All], I was like I would never need that (P22).”

When the purpose was non-specific, in which participants were either randomly browsing content or checking for apps ad-hoc, such as browsing random apps or checking apps when a notification arises, 13 out of the 24 participants (54.17%) voted for the *None* interface, 8 (33.33%) voted for *Some*, and 2 participants (8.33%) chose the *All* interface. Participants who liked *None* preferred its *controllability* and *on-demand* characteristic: “In situations that I need random stuff, I would like to be in a little bit control (P8);” “when I need random apps, I could use None and open when I need them (P13).” Two participants voted for *All*: “If I am just randomly browsing, it does not bother me, I can just have all the apps there (P15).” However, the majority of participants think *All* is still too *cluttered* and *irrelevant*,...
5.3. Evaluating Realistic Transitions In-Situ

similar to previously listed comments. Participants voted for Some because of its potential to fulfill unspecific information needs: “sometimes the apps I need were already opened by the system, which I found useful (P7);” “you can potentially have the apps ready without the risk of being cluttered (P14);” and “…if the prediction is really good [in figuring out what I will need], I would use Some (P23).”

Desired features

In the interview, participants also touched upon some functions they would like to see included in the three interfaces.

Customizing the Some apps. Seven participants (29.17%) mentioned that it would be great if they could be involved in defining what apps will show up when they enter a new room. Participants commented “The Some interface may be really good for someone who would like to add the things to different locations so it is more personalized for them (P3);” and “the best scenario is you preselect what app I want opened in Some. It would be better if I can define what apps are opened in each space (P4);” and “I love the convenience of the Some, especially if I could personally set what I wanted in what room, would make things really great (P8).”

None with pinning. Four participants (16.67%) mentioned that it would be cool to add a Pin function to the None interface. When they leave the current space, this feature would allow pinned apps to remain open in the new room: “When I need to carry the note app from one place to another, I would prefer the list to follow me around all the time, like if I can pin the app, so the app just follow me around wherever I go (P24);” and “what if the None has a default follow feature, like when I open this app, I want this app to be opened when I move around, that could be helpful so I do not have to reopen the stuff [apps] I used
Dimming all windows. Two participants (8.33%) mentioned that they would like a function to completely wipe out the AR applications temporarily if the current task requires heavy attention on the physical world. “If there is a toggle button, where if I wanted, I could use a toggle a switch to make everything disappear just temporarily. When I turn it back, all apps comes back (P24),” and “I think it would be helpful to have a gesture to wipe things away temporarily, and bring things back later. When I want to search for something, I could easily wipe things away real quick, to find the site, and bring the apps back. With this, the Some display would be even better (P10).”

5.3.7 Discussion & Design Implications

Impacts of the contextual factors

Looking back at our hypothesized trade-offs, our hypothesis about None and Some being better options than All when the real world is prioritized was supported by our study. Participants favored the clear visibility of the physical environment that None offers. Some was also favored by over half of the participants, who did not consider it distracting even when attention to the physical surroundings was required. In contrast, no participants voted for All, criticizing its clutter and occlusion. Participants had to rotate their heads to look past the AR windows in order to see the real world, which was also shown by the significantly higher head rotation produced by the All interface. When the real-world is not prioritized, our hypothesis partially aligns with the findings, since Some was preferred over None.

When it comes to workload, our hypothesis was partially supported: Some was the most favored interface in high-workload situations. We hypothesized that All would also be favorable since users could avoid the extra effort to open apps manually. To our surprise,
participants preferred *None* rather than *All*. The potential distractions of *All* and the need
to manually close the apps when they got in the way outweighed the benefits of having the
apps automatically opened in demanding situations. In *None*, participants liked the one-
handed gaze-triggered menu design, thought it got in the way less, and that it felt similar
to how they access information with phones or laptops. This demonstrates the importance
of optimizing the manual approach; our redesigned menu made *None* more usable than it
was in Study 1. In low-workload situations, participants favored *Some*, which contradicted
our hypothesis about all three interfaces being equally usable. The suggested apps provided
by *Some* provided participants with the convenience of leveraging already-available options,
while not being overwhelmed by too much information.

Finally, when it comes to specific vs. non-specific needs, our results supported our hypothesis
by showing that *Some* was the most favored interface in situations when the users need
specific applications to assist their tasks. As reflected in the ratings, the automation in *Some*
was perceived to be more accurate than *All*. Participants considered it natural to have task-
relevant applications automatically opened. *All* would have too many irrelevant applications
opened in the space, which was considered distracting and unhelpful, also demonstrated by
its lower perceived accuracy level than *Some*. When it comes to non-specific uses, our
hypothesis of *None* being the preferred interface was supported, but not *All*. *None* was
favored for its controllability and on-demand nature, which makes it viable for handling the
ad-hoc information needs of the users.

**Some as the sweet spot between None and All**

Our findings in study 2 strongly indicate the value of *Some* in supporting everyday users
to access information after spatial transitions. It was the most preferred interface, received
higher usability ratings, had a similar agency rating as the manual condition, and had a lower
task completion time (though not significantly lower). In general, participants’ comments suggested that *Some* is a good balance between *None* and *All*, and nine participants (37.5%) stated this explicitly. It offered more convenient access to relevant applications as compared to the fully-manual approach of *None*, while having less risk of distraction, occlusion, and showing irrelevant applications to the task and spaces as compared to *All*.

Additionally, participants were positive about the semantic connections between the automated applications and the physical objects in *Some*. Participants commented “*Having the recipe app automatically opened near my fridge, plant app opened on my plant, and the email app opened in my office make lots of sense* (P25);” “*A lot of people have cook books, recipes on their refrigerator, so it make sense to have the [recipe] application there … it makes more sense to attach the app to the object* (P3);”. Having the AR apps appearing right next to the physical referent to assist the interaction felt natural and useful. Participants also mentioned that such connections could work as a reminder: “*The plant app is nice to have open near my plant, because it is easy to forget that the plant needs watering* (P26).”

In contrast, participants criticized it when the AR apps appeared in ways that did not show connections to the physical space. For example, P10 commented on the *All* condition: “*The plant app is opened in the Office, but the plant is not actually there, which I felt did not make much sense* (P10).” Interestingly, participants also mentioned how the “feel” of the physical space needs to be considered: “*In the kitchen, you are already up and ... walking around. So it makes sense to raise your hand [for the menu], like the none or the some interface. Like in the office, you would want to be more ... cognitive. In the living room, when you are kinda lazy, I would prefer to have the apps already opened for you, like the all and some* (P14);” and “*Having message and mails up in the office makes much sense. But having them in the living room, which is supposed to be a relaxing space, it is not supposed to have email there, you know, work-related stuff* (P8).” This is related to the notion of relevancy. The apps need
to be automated or provided in ways that respect the users and their environment, taking into consideration the potential types of activity being done in the particular space. This was well-achieved in the Some interface in study 2, in that only the apps relevant to the space would appear, and the tasks that took place were relevant to the space in the design of the scenarios.

5.3.8 Design Recommendations

Based on our findings in both studies, we distill the following design recommendations:

- *Choices of the interfaces.* Use All when the number of apps is relatively low and information access time is the most critical component of effectiveness. Use None when information access time is less important, but strong attention is required on the real world environment. Use Some when the number of apps is relatively large, when users need specific task-related applications, and when the predicted accuracy level of the automation is satisfactory.

We recommend considering the following implications for deploying world-anchored transition-aware interface solutions for everyday AR uses:

- *Valuing real-world visibility.* Even in cases where less attention was demanded on the physical environment, most participants still preferred having good visibility and awareness of their physical surroundings. World-anchored transition-aware interfaces should not cover up too much of the physical space. If the occlusion is unavoidable, it is important to make the real world easily accessible to the users, maybe through a feature that temporarily makes all AR content invisible.
• Relevancy over Availability. In general, the relevancy of the AR content to the task and space is valued more than the simple availability of the AR content. Instead of having more applications all readily available to eliminate the need for manual input, it is preferable to have a more relevant subset opened automatically, and then empower users to pin critical apps that should be opened everywhere, and to open the less-relevant ones on-demand manually with a menu interface.

• Respecting the physical space. While automating which apps to open, take into consideration the space that the user is in and its characteristics. Recommending work-related applications in a casual environment, or suggesting apps that are less relevant to the themes of the tasks to be done in the environment, would not be desirable.

• Leveraging semantics. When spawning or placing AR applications automatically, it is helpful to leverage existing physical objects in the space and place the AR app in proximity to the objects that are semantically close. Examples include the recipe app near the fridge, the email app near the laptop, or the plant app on top of actual plants.

While deploying the Some interface specifically, we recommend considering the following implications:

• Awareness of already-opened apps. For inexperienced users who are less familiar with what apps are opened for every space, we need to design ways to assist their awareness of the already-opened apps. Having a list of icons on the top of the display indicating the already-opened apps, as used in the second study, was demonstrated to be helpful.

• Optimizing the menu design. The design of the menu needs to be optimized. As such, it would not be troublesome if users need to manually open apps that were not predicted by the system (i.e., correct system errors). We suggest using a one-handed or hands-free interaction design, taking into account that the users’ hands are often
occupied by physical activities.

5.4 Limitations & Future Work

Our studies have some limitations. First, study 1 was conducted in a controlled simulation of transitional activities and was less ecologically valid. This was resolved in study 2. Second, in study 2, we used a public space rather than users’ private spaces. The users’ familiarity with the spaces may play a role in their perceptions of the interfaces. Third, although given time to become familiar with the layout and content of the AR applications in each space, the placements of the apps were not defined by the participants themselves. This could have caused negative impacts on the Some and All interfaces. Fourth, the scenarios, although composed of easy-to-understand tasks that commonly happen in everyday life, were artificial in study 2. Future studies could explore self-triggered uses of the interfaces in uncontrolled everyday scenarios. Fifth, study 2 was conducted in the context of a home environment, which was quiet and indoors. Future research could study how user preferences change in more crowded, open environments. Sixth, in both studies, we maintained spatial consistency of AR windows across multiple physical environments. This could be challenging to achieve if these environments have different spatial features or layouts (e.g., a plant is in the left corner of room A but in the right corner of room B; room A has a desk but room B does not). The locations, sizes, and layouts of certain AR applications may have to be adapted, which may disrupt the established spatial consistency and mental model of the users. Future research could study such trade-offs, how to maintain the balance between spatial consistency and semantic coupling between AR content and its physical surroundings across multiple environments [50]. Last, previous research has demonstrated that allowing users to switch between predefined sets of layouts could be useful in assisting computational tasks [72].
Future work could explore such possibility to enable a more adaptive and controllable layout adaptation experience.

5.5 Conclusions

Current AR systems offer limited solutions for users to carry previously placed AR content to new spaces for continuous access. In this research, we explored the design of transition-aware interfaces in head-worn AR to assist users in accessing information when moving around multiple physical locations. We explored three interface options: (1) None, in which users rely on a menu to manually open an application while entering a new location, representing existing solutions; (2) Some, in which the system automatically suggests a subset of the total applications; and (3) All, in which the system automatically opens all applications when the user enters a new space. Through a modeling study, we identified how the number of existing applications influences the user performance on retrieving a target application. Utilizing the learnings from study 1, we further improved the design of these interfaces and compared them in an ecologically-valid simulation of everyday transition scenarios. Our results provide valuable lessons on how to design interfaces that are scalable and transition-aware, how to take into consideration different contextual factors, and how to improve the user experience in future everyday AR glasses.
Chapter 6

Real-world Deployments of Glanceable AR

6.1 Overview

In 2016, Grubert et al. proposed “Pervasive AR,” which refers to “a continuous, omnipresent, and universal augmented interface to information in the physical world” [65]. With augmented virtual information being pervasive and always available, certain issues could be introduced in everyday scenarios. The virtual information might become overwhelming, occlude real-world objects of importance, distract or disrupt users from their real-world tasks, and further lead to privacy issues when interactions are visible by co-present others. As pointed out by Grubert et al., when AR experiences become omnipresent, we need to design “appropriate information display and interaction, which is unobtrusive, not distracting, and is relevant and safe to use [65].”

Glanceable AR, proposed in chapter 3, is an interaction paradigm that proposes to address these issues. Glanceable AR allows information acquisition through quick glances at the periphery of the visual field [102]. The paradigm builds upon existing research on glanceable displays [18, 25, 114, 119] and peripheral awareness [71, 164], which are proven to be viable for quick access and monitoring of secondary information. It allows unobtrusive display of information together with quick intake, which could be a promising way to achieve Pervasive
AR in everyday scenarios.

To demonstrate the potential of Glanceable AR, two major challenges need to be addressed. First, we must address the design challenge of creating practical Glanceable AR widgets that are combined into a usable system for real-world tasks. Existing work on Pervasive and Glanceable AR has produced small proof-of-concept implementations, but has stopped short of implementing a practical system that can be actually used in everyday situations. Second, we need to address the validation challenge by demonstrating that such a system can provide a high-quality user experience, a high level of user acceptance, and a high level of usefulness for everyday tasks in real-world scenarios. Conventional AR user studies typically happen in controlled lab settings; there are a very limited number of studies that have explored in-the-wild use of AR interfaces, especially with unsupervised users in real-world scenarios.

In this chapter, we address these challenges. We describe two Glanceable AR prototypes and two studies. First, we implemented a video prototype and distributed it broadly, along with a survey, to understand user perceptions and acceptance of Glanceable AR. Second, we implemented a working Glanceable AR system on the Magic Leap One AR headset. The application is able to display time, weather, and news, along with users’ personal calendar events, incoming email, to-do list, and fitness data as glanceable widgets. Using this prototype, we conducted an in-the-wild study in which participants freely used the application in their everyday lives for at least 2.5 hours.

The contributions include: (1) explorations of user acceptance and perceptions of applying AR HWDs with a glanceable interface for everyday tasks; (2) design and implementation of a working Glanceable AR system for everyday uses; and (3) evaluations of the interface through an unsupervised in-the-wild study that explored authentic everyday use cases.

The section 6.2 and section 6.3 in this chapter has been published at IEEE VR 2021.
6.2 Surveying Public Opinions About a Glanceable AR Video Prototype

6.2.1 The video prototype

To convey our design ideas in distributable form, we developed a video prototype\footnote{https://youtu.be/qsDVDP7wA-E} that illustrates some envisioned scenarios of using Glanceable AR in common everyday situations. We showcase three scenarios in the prototype (Figure 6.2): (1) working in front of a desktop computer; (2) taking a walk outdoors; and (3) cooking according to a recipe. The reasons

\begin{itemize}
\end{itemize}
we chose these scenarios are two-fold: (1) they are common and are good representatives of sitting, standing, walking conditions requiring and not requiring hands; (2) they all prioritize focus on the real-world environment or objects, in which accessing digital information on separate devices could be intrusive and distracting. The video prototype walks through the three scenarios one by one with an actor wearing a Magic Leap One AR headset. The prototype uses a Wizard-of-Oz approach [41] for interaction. In the following section, we will explain the core elements we embedded in the prototype, and the design rationale behind them.
6.2. Design rationale

Glanceability

As its name implies, glanceability is a major design principle of Glanceable AR [102]. 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 [109].

Widgets are common in mobile phone and desktop computer interfaces [154]. They continuously run in the background [19], and allow users to quickly glance at information without opening an app. In everyday AR scenarios, widgets could be a feasible form of presenting multiple pieces of information simultaneously and continuously to users. They are compact so they take a small proportion of the field of view, and they are easy to access and understand with a quick glance. Grubert et al. found that AR widgets could be feasible while interacting with physical displays [64]. Based on our review, there is limited research that explores the potential integration of widgets into AR systems for general-purpose information display. As such, we showed nine widgets in the video prototype that people use frequently, including Weather, Timer, News, Calendar, To-do List, Music, Fitness, Recipe and Email. To avoid distracting or disrupting users’ tasks in the real world, following the design principles of Glanceable AR, these widgets are positioned at the periphery to be unobtrusive, but can be accessed via a glance whenever needed.

Interface design

Four types of interface design choices were shown in the video prototype: gaze, follow mode, voice&hand gestures, and notifications.
**Gaze:** To allow for an adaptive information display, we demonstrated a gaze-contingent metaphor \[62\]. The amount of information in the widgets is dynamically managed depending on the direction the user is looking. The widgets are normally displayed in low detail to minimize the space they occupy, however, once gazed at, they expand and show more information. For example, the weather widget shows current weather and temperature by default, but it expands to show a weather forecast for the next few hours when looked at. We demonstrated the gaze interactions with the actor gazing at the email, calendar, fitness and weather widgets.

**Follow Mode:** The widgets incorporate two modes: stationary mode and follow mode. Adopting the design of existing glanceable and peripheral displays, the widgets are in stationary mode by default and are world-fixed. Following the “Fixed to Users” design principle of Glanceable AR (see subsection 3.2.2), the widgets can enter follow mode to be body-fixed and follow the users around to enable information access on the go \[102\]. For stationary scenarios such as working in front of a computer monitor, the widgets are placed at the edge of the monitor to avoid distractions. For moving scenarios such as cooking and exercising, the head-glance (HG) interface was applied \[102\], in which the widgets are placed outside the periphery of the user to ensure clear forward vision, but can be accessed quickly by turning the head. In the video, we demonstrated both stationary (see Figure 6.2(a)) and mobile use cases (see Figure 6.2(b)).

**Voice & Hand Gestures:** We illustrated two other input modalities in the video prototype. Voice could be useful when hands are occupied for other tasks, while hand gestures could be useful when voice is not convenient in a given scenario. The actor uses voice to set a timer, play music, and trigger the widgets to follow him. He uses hand gestures to browse the music library while exercising outside.
6.2. Surveying Public Opinions About a Glanceable AR Video Prototype

Notifications: Notification system are crucial for personal computing devices. Existing mobile devices mainly apply auditory and tactile cues to notify users of certain events [68]. Unlike traditional mobile devices in which digital content is constrained to the physical device, in AR content can be displayed around the user in the 3D environment. As such, notifications in AR HWDs should not only deliver information, but also the location of the virtual content so users can be spatially aware of the piece of content that is drawing attention. Costanza et al. found that changing the brightness and speed of visual stimuli at the periphery could be an effective strategy to deliver notifications in HWDs [38]. Similarly, in our prototype, we display the icon of a widget and an arrow in the periphery and alter the brightness to notify users. The arrow points towards the widget that is delivering the notification so that users can easily find the widget in the physical space.

6.2.3 Research Goals

In the first study, our goal was to obtain feedback from a broad sample of the general population about the Glanceable AR design shown in our video prototype, and to understand their perceptions of using it in different scenarios, compared to the ways they accomplish those tasks now. We aimed to answer the following four research questions:

- **RQ6.A.** How do users perceive the quality of user experience of the Glanceable AR prototype in everyday use cases?

- **RQ6.B.** What are the perceived benefits/drawbacks of using AR systems like this?

- **RQ6.C.** What are the scenarios that users perceive as beneficial when using our Glanceable AR approach?

- **RQ6.D.** What are the pros and cons that users perceive for using Glanceable AR
as compared to existing personal computing devices (e.g., smartphones, tablets, and smartwatches)?

### 6.2.4 Experiment

We developed an online survey to gather data about people’s perceptions of the video prototype. The study, which was approved by our university ethics board, was voluntary/unpaid and consisted of six steps. Participants were first asked to read and agree to the consent information presented at the beginning of the survey. Second, they were asked to provide their background information. Third, they were asked to watch the video prototype. After they finished watching, fourth, they were asked to fill out the System Usability Scale (SUS) and a shortened version of the User Experience Questionnaires (UEQ) on overall and individual features of the interface [22, 140]. This allows us to assess the overall usability as well as dive down to each usability subcategory. Although such questionnaires are typically used following a participant’s actual interaction with an interface, we expected that they would still give us information about participants’ perceptions of expected usability and user experience (UX) after watching the video. We restrict our claims to such perceptions and do not make any claims about actual usability and UX from this questionnaire data. Fifth, they were asked to imagine using the Glanceable AR interface in real life, and compare it with existing devices that they frequently use in terms of *ease of access* and *distraction* for six tasks listed below:

- **Task 1:** Check information while stationary.

- **Task 2:** Check information while moving.

- **Task 3:** Check information while hands are occupied.
6.2. Surveying Public Opinions About a Glanceable AR Video Prototype

- **Task 4:** Interact with widgets while stationary.

- **Task 5:** Interact with widgets while moving

- **Task 6:** Interact with widgets while hands are occupied

Finally, they were asked to give comments on what they liked or disliked, and how the interface is better or worse as compared to existing devices they normally use. At the conclusion of the survey, they were asked to leave their email address if they would like to be contacted for an additional 20-minute interview. In the interview, we asked them to explain their comments in the questionnaire and provide further suggestions.

In order to evaluate only the Glanceable AR interface design, we asked participants to ignore the form factor of the Magic Leap headset in their SUS and UEQ responses. The survey took around 30 minutes to complete. It was distributed across multiple online communities, including Reddit, Twitter, Linkedin, Facebook and university mailing lists.

### 6.2.5 Participants

We received 123 responses for our survey, among which 60 responses (48.78%) were abandoned due to incompleteness and low quality. 63 complete responses were recorded successfully in the end. The respondents (14 females, 49 males) were between 20 and 69 years old ($M = 32.90, SD = 11.25$). Although gender balance was desired, we did not fully achieve this since participants self-selected to participate. Thirty-one of them self-identified as experienced with AR. Respondents came from various backgrounds, including college students (19), lecturers/researchers/professors/postdocs (11), developers/engineers/designers (9) and consultants/advisors/analysts (4). Participants were from all over the world; the most frequent countries of residence were USA (24), China (8), Italy (5), Germany (4) and France.
CHAPTER 6. REAL-WORLD DEPLOYMENTS OF GLANCEABLE AR

Almost all (62) of the participants said they used mobile phones everyday. Participants also used personal computers (60), tablets (34), smartwatches (23), and virtual assistants (21). Eight respondents participated in the additional interview.

6.2.6 Results

We conducted a series of analyses on our results. The Shapiro-Wilk test was applied to test the normality of the quantitative data we obtained. Non-parametric tests (i.e., the Mann-Whitney U test and the Wilcoxon signed-rank test) were conducted for non-normally distributed data. The $r$ statistic was reported as a measurement of effect size $[122, 135]$. According to Cohen’s classification $[37]$, $0.1 \leq r < 0.3$, $0.3 \leq r < 0.5$, and $r \geq 0.5$ would be considered as small, medium and large effects respectively. We used an $\alpha$ level of 0.05 in all significance tests. In the results figures, pairs that are significantly different are marked with * when $p \leq .05$, ** when $p \leq .01$, and *** when $p \leq .001$.

Perceived user experience

The SUS scores of the video prototype ranged from 15.00 to 95.00 with a mean score of 67.14 (SD = 16.53), which is between ‘OK’ and ‘Good’ $[11]$. The Mann-Whitney U test yielded no difference for SUS score between experienced and inexperienced AR respondents ($U = -485.5, Z = .145, p = .885$). Figure 6.3 (a) shows the results of the overall UEQ together with UEQ ratings (for both pragmatic and hedonic scales) for the following six features: (1) glance-stationary: glancing at the AR widgets while stationary to acquire information; (2) glance-mobile: glancing at the AR widgets while moving to acquire information; (3) gaze: gazing at the widget for higher detail; (4) follow mode: attaching widgets to the user for mobile use; (5) notification: delivering notifications through blinking icons at the
periphery pointing to the source and (6) Voice&Gesture: using voice input or hand gestures to control the widgets being gazed at. All of the average ratings were greater than 0.8, which is considered positive. When asked if they would want to use the interface on a daily basis if the form factor of the display was similar to a pair of eyeglasses, 36 participants gave a positive response (57.14%), 18 respondents (28.57%) selected unsure, and 9 participants (14.29%) gave a negative response.

Figure 6.3 (b) shows the mean ratings for Task 1 - Task 6 in terms of ease of access and distraction. Wilcoxon signed-rank tests showed that participants rated the video prototype as being significantly easier to acquire information as compared to the existing devices they normally use for Task 2 ($Z = 3.335, p = .001, r = .297$), Task 3 ($Z = 5.796, p < .001, r = .516$), and Task 6 ($Z = 4.933, p < .001, r = .439$). The approaches shown in the video prototype were also rated as significantly less distracting compared to existing devices for Task 3 ($Z = 3.428, p = .001, r = .305$) and Task 6 ($Z = 3.013, p = .003, r = .268$). The Mann-Whitney U test shows that for Task 1, respondents inexperienced with AR considered the video prototype as easier for accessing information, as compared to respondents experienced with AR ($U = 337.5, Z = 2.262, p = .024, r = .285$).

Qualitative Results

We used the qualitative analysis software AQUAD 8 to perform qualitative coding on approximately 391 lines of text from comments in the surveys and interview scripts [137]. We categorized and grouped the codes, and listed the most frequently-mentioned ones together with the frequency of occurrence in Table 6.1.

As for perceived benefits of the video prototype compared to existing devices, codes such as hands-free, convenient, awareness of information, non-intrusive, always accessible, and real-
Figure 6.3: (a) UEQ results for both overall and separate features shown in the video prototype; (b) ratings of ease of access and distraction for six tasks (higher ratings mean easier access and lower distraction) (±S.E.).

time data appear most frequently. Participants commented: Information is always accessible; I like not being distracted; you can still maintain awareness of what is going on around you, and it makes it easier to access information when your hands are occupied.

The most commonly mentioned potential drawback of the interface, because the actor was wearing the Magic Leap One headset, was the form factor of the display. A participant commented glasses must be really small and stylish before being considered viable. Participants also disliked voice interactions shown in the prototype as it could disturb others in a shared space, which is consistent with existing literature. Other than form factor and voice, distracting, overwhelming, invasive, occluding the real-world were also mentioned by several
Table 6.1: Most frequently occurring codes and their frequencies in four categories: benefits of the interface, drawbacks of the interface, wish-to-have features, and other scenarios in which the interface could be useful.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Drawbacks</th>
<th>Wish-to-haves</th>
<th>Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hands-Free(30)</td>
<td>Form factor(28)</td>
<td>Customization(16)</td>
<td>Navigation(15)</td>
</tr>
<tr>
<td>Convenient(27)</td>
<td>Voice input(20)</td>
<td>Phone call(9)</td>
<td>Shopping(9)</td>
</tr>
<tr>
<td>Awareness(19)</td>
<td>Distracting(19)</td>
<td>Messaging(7)</td>
<td>Assembling(5)</td>
</tr>
<tr>
<td>Non-intrusive(13)</td>
<td>Overwhelming(18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Always accessible(8)</td>
<td>Invasive(12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real-time Data(5)</td>
<td>Occlusion(8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>Addiction(5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

participants. They commented the content could still block my vision when I’m looking to the side (at the real world), I think the apps would actually get in the way too often, and I worry about the increase in the number of distractions it would bring, and how it could potentially take away my vision. Five participants felt worried that the interface could cause addiction to virtual content due to its ease of access and always-on nature, which could limit their interactions with the real world: I already feel like I spend too much time checking emails ...

I’m afraid that wearing such a device might lead to an over stimulation from the technology, and I think having stuff floating around all the time might be distracting and make me even more addicted to the internet.

For wish-to-have functions that were not demonstrated in the video prototype, customizing the virtual content (widgets) appeared the most, including location in space and notification levels. Phone calls, messaging and watching videos were also brought out by respondents. As for other scenarios that users would like to use the interface in, navigation was mentioned by the most participants. Fourteen participants also mentioned having a shopping list/instruction tutorials at the periphery while hands are busy during shopping/assembling furniture could be potential use cases.
6.2.7 Discussion

In this section, we will summarize and discuss our findings based on the RQs.

**RQ6.A.** The pragmatic and hedonic UEQ ratings were positive (>0.8) both overall and for separate features, among which the gaze and follow mode features obtained especially high scores, indicating that people expect them to be practical and pleasant to use. Over half of the respondents affirmed that they would like to use the interface everyday, which provides further evidence for the potential of Glanceable AR for everyday use cases.

**RQ6.B.** It was interesting to us that respondents held contradictory opinions on the benefits and drawbacks of the interface. Some of them considered it non-intrusive and convenient, while others considered the interface distracting and overwhelming. One potential reason for this could be the limited expressivity of the video prototype. In the interview, two participants commented: *it is hard to tell how distracting notifications are just from the videos, and I think the apps would actually get in the way too often, but I’d have to test out the system on my own to see if my hypothesis is correct.* This could also be reflected by the wide range of SUS scores, and that 28% of participants felt unsure about whether they would want to use the system frequently. Without actually experiencing the interface, the opinions of participants could be incomplete or even biased by their own understanding of the video. As such, although the prototype obtained a positive response in general, evaluations of actual use are needed to validate the user experience of Glanceable AR.

**RQ6.C & RQ6.D:** As shown in Table 6.1, the most compelling scenarios to participants were those where the hands are occupied with tasks in the real world. The ratings on the tasks also show that compared to existing devices, participants considered the Glanceable AR interfaces significantly more convenient and less distracting in scenarios when hands are occupied. As shown in Figure 6.3(b), checking information while moving with Glanceable
AR was also considered more convenient as compared to existing devices.

To summarize the first study, we conveyed the idea of Glanceable AR to a sample of participants of varied ages and backgrounds through a video prototype and obtained their feedback. Our results show that participants feel relatively positive about Glanceable AR in general, but some improvements should be made to the prototype, and issues such as distraction and occlusion cannot be fully understood with the video. The usability and acceptance of the Glanceable AR interfaces need to be explored through direct user experience.

6.3 Capstone Implementation I: Initial Perceptions in Authentic Everyday Uses

6.3.1 Working Prototype

To explore authentic uses of Glanceable AR interfaces, we implemented a working prototype on the Magic Leap One AR headset. Similar to the video prototype, the working system incorporated seven widgets: Email, Calendar, Fitness, Tasks, News, Weather, and Clock. Using this prototype, we conducted a second study to explore the user acceptance, requirements and experience of using a Glanceable AR interface in the wild. Compared to most AR user studies, this study was unique in several respects: (1) participants used the prototype for a minimum of 2.5 hours in authentic everyday use cases; (2) the sessions were unsupervised with a minimal level of outside control; and (3) users’ own data were integrated into the working prototype to allow practical uses.
Personal & Real-world Data

All widgets were linked with real-world or personal data sources. The Gmail, Calendar, Fitness, and Tasks widgets were linked with users’ personal Google account. The Gmail widget displayed the latest or incoming email from users’ Gmail inbox. The calendar widget displayed three calendar events, the current, previous and next event from users’ Google calendar. The Fitness widget displayed personal data from Google Fit, which includes calorie consumption, daily walked steps, and active minutes, and could be linked to sensors in personal fitness tracking devices. The Tasks widget displayed to-do lists the user entered in a google sheet. To implement these features, we asked participants to give permission for our application to access data associated with their Google account, and compiled a personal version of the application for each participant with an access token generated by Google. We did not collect participants’ account names or passwords, nor did we have access to log in to their Google accounts. We deleted the access tokens as soon as we finished deployments, and we deleted the personalized versions of our application as soon as participants finished the study. Participants could remove the access they granted to the AR application to access their google information at any time in their Google account settings.

The News, Weather and Clock widgets displayed real-world information. For example, the News widget displayed random news headlines from users’ region. The Weather widget displayed real-world weather conditions, temperature and a five-hour forecast. The clock widget displayed current local time, and also provide alarm and countdown timer functionality.

Notification

Given the contradictory feedback about notifications from the first study, we decide to keep the feature in the working prototype to study it further. When a notification happened,
the icon of the notifying widget appeared and blinked at the edge of the FoV of the display. The icon pointed towards the location of the widget so users could easily locate it. We integrated notification systems in all seven widgets. A notification was delivered for the following conditions: (1) Gmail: when a new email arrived in the inbox; (2) Calendar: when an event was approaching; (3) Fitness: when daily goals were completed; (4) Tasks: when the deadline of a task was approaching; (5) News: when the widget obtained a new news story; (6) Weather: when rain was forecast in the next hour; and (7) Clock: When a timer expired or alarm was triggered.

6.3.2 Improvements

Gaze: Given the positive feedback on gaze from the first study, we kept the gaze-contingent metaphor for the working prototype. For all the widgets, gazing at the widget would either expand the widget to show more detail, or trigger some interaction. For example, the weather widget displayed current temperature and weather condition by default, but expanded with weather forecasts when gazed at. We added a one-second dwell time to avoid the “Midas touch effect [82].” A progress circle would appear on the side of the widget to represent the dwell time (see Figure 6.4). As such, the user could always perform a quick glance to look at the low-detail information, and gaze for a longer time only if they needed more information from a widget.

Follow Mode with FG: In Section subsection 3.2.3, we introduced the head glance interface as the primary method of on-the-go information access. However, as pointed out by respondents in the survey, even if the widgets are placed outside the periphery to be glanced at via head rotation, occlusion issues could still happen when users intend to turn their heads to look at the real world instead of the widgets. As such, we employed the technique
CHAPTER 6. REAL-WORLD DEPLOYMENTS OF GLANCEABLE AR

that we developed in subsection 4.3.2 called Fixation Glance (FG) (see Figure 6.5). In FG, widgets were represented as small visual targets, and the full widget only appeared if user was looking in the direction and at the depth of the target. The Magic Leap’s dual eye trackers were used to determine both gaze direction and vergence depth. As such, if the system detected that the user was looking behind (or in front of) the target, the widget would not appear, to avoid occluding people’s vision. Existing research has explored applying gaze depth for virtual object selection and manipulations [124]. Hirzle et al. explored applications of gaze depth to enable x-ray vision in HMDs [73]. To our knowledge, there is little research on applying gaze depth for information acquisition. We see FG as a promising strategy to enable a more unobtrusive information display compared to HG. In the second study, we integrated FG into the working prototype.

Customization: As the top wish-to-have function in the list, we decided to include customization in our prototype. Survey respondents wanted customization of widget positions and level of notification. As such, in the working prototype, we allowed users to place and scale the widgets freely (see Figure 6.6 (a)). For follow mode, users could specify which location in the periphery they wanted to assign each widget. We also allowed users to customize notifications in a separate menu (see Figure 6.6 (b)), for example, how far in advance they would like a calendar event to give a notification. Also, users could mute any of the
6.3. **Capstone Implementation I: Initial Perceptions in Authentic Everyday Uses**

Figure 6.5: An illustration of the Fixation Glance (FG) technique: (a) widgets are represented as small targets to avoid occluding users’ view when the user is looking at the real-world environment behind the target; (b) when the user converges their gaze at the depth of the target, the widget expands and appears.

Figure 6.6: (a) Using raycasting to select and move widgets; (b) Adjusting news category and notification levels in the main menu.

seven widgets if they did not want to receive notifications from them, and could customize the category of news items. Users could freely switch between HG and FG for mobile use. To ensure robust control, we did not use voice or hand gestures. All customizations were achieved through the Magic Leap One controller via ray-casting [20]. Users kept the controller in a 3D-printed case attached to a belt; this allowed us to track torso orientation to enable head glancing, and also kept the controller conveniently available for customization when needed.
6.3.3 Research Goals

In the second study, our goal was to evaluate Glanceable AR in authentic everyday scenarios in an unsupervised manner to understand user acceptance and perceptions. We aimed to answer the following three research questions:

- **RQ6.E.** How do users perceive the quality of user experience of the Glanceable AR system in authentic everyday use cases as compared to existing personal computing devices (e.g., smartphones, tablets, and smartwatches)?

- **RQ6.F.** What are the common user behaviors and patterns when using the Glanceable AR system?

- **RQ6.G.** When using the Glanceable AR system, for what scenarios is it considered beneficial or unfavorable?

6.3.4 Participants

We recruited six participants (2F/4M) between 23 and 29 years old (M=25.33, SD=2.58). Participants were all college students. Three participants self-identified as experienced with AR. All participants used Google services (including Gmail and Google calendar) frequently for everyday work. All participants reported that they used mobile phones and PCs every day, while some used virtual assistants (4), tablets (2) and smartwatches (1). Four participants participated in the study locally on campus, and two participants participated in the study remotely. For local participants, a Magic Leap One AR headset was provided for them to take home for the duration of the study. Remote participants needed to have access to a Magic Leap One headset to be able to participate.
6.3.5 Experiment Procedures

The experiment, which was approved by our university ethics board, was divided into four phases. In the first phase, participants were asked to read and sign the digital consent form remotely. In the second phase, an online tutorial session was provided to participants to walk through the hardware, calibration processes, and the Glanceable AR interface. The length of this phase ranged from 80 to 120 minutes, depending on how experienced participants were with the hardware. In the next phase, participants were asked to freely use the interface in everyday scenarios for at least five sessions of at least 30 minutes over the course of three days. We chose 30 minutes empirically because it was a suitable time to wear the Magic Leap 1 AR headset without feeling uncomfortable or visual fatigue, so participants could ignore hardware constraints as much as possible while offering feedback. Then participants were asked to fill out a diary survey immediately after completing each session. In the diary, we asked about the time period, scenarios of use, layouts of widgets, and perceived user experience in that session. In the fourth phase, participants were asked to complete a post-study survey (SUS and full version UEQ), and participate in a 40-minute final interview, which was audio-recorded. After all study sessions were complete, participants were compensated with $70 USD.

6.3.6 Results

Usage sessions and scenarios

We lost three sessions of log data accidentally during data transfer. Thus, 27 sessions were recorded successfully, yielding a total of 936.61 minutes (15.61 hours) usage time. Various scenarios of use were reported by participants in the diaries, including working in front of a computer (reported in 16 sessions); cooking/eating (7 sessions); watching TV (5 sessions);
playing games (3 sessions); attending a remote classes or meeting (3 sessions); talking on the phone (2 sessions); cleaning/doing laundry (1 session) and brainstorming on a whiteboard (1 session).

**Overall user experience**

The SUS scores ranged from 62.5 to 97.5, with a mean score of 90 from AR experts and 73.33 from non-experts, yielding an overall mean score of 81.67 (SD = 15.54), which is very positive. Figure 6.7(a) shows the mean ratings for the six UEQ subcategories. Based on a benchmark published in 2017, all six categories obtained better results than 90% of existing studies[139].

**Arrangement strategies**

Based on our playback of data from the logs of all 27 sessions, we observed three patterns that participants used to arrange the widgets when they were stationary.

The first common strategy we observed was that participants liked to arrange the widgets around a physical object that was their primary focus. For example, for the 22 sessions that involved participants working in front of a desktop monitor, attending a meeting/class, or watching TV, 16 (72.73%) of the sessions demonstrated strong patterns of arranging the widgets on the edge of the computer monitor, laptop screen, or TV, (as in Figure 2.1 (a)). Participants commented that: *I like putting the widgets close to my monitor. I just look slightly beyond the monitor, and they are there; and having them around my monitor makes me feel my screen is extended, I do not need to have email window open on my monitor, which saves my screen space.*

The second common strategy we observed was clustering and placing widgets depending on
the perceived frequency of use in the session. In the playback of 11 sessions (40.74%), we observed that participants grouped the widgets. A participant commented *it’s like when you organize your desk, you move some things closer to you and some things further away depending on how important they are to your task.*

The third strategy, which was interesting but uncommon, was attaching widgets to different locations around the house. We did not collect any information about participants’ environment due to privacy concerns. However, a combination of the interview, the diary entries for
each session, and the playback allowed us to confirm that widgets were placed in different locations/rooms in a house by two out of the participants. One participant commented: *I tried to position the widgets as if they were at different locations in my house, and I really liked it;* and the other mentioned *I place the weather widget at my door near my keys, so as in situations where I will leave my house, I can quick glance and know if it is going to rain.*

**Interview comments**

**Perceived user experience:** The interface received positive feedback from all participants. Participants commented that: *I liked it a lot! I think it is really convenient, if it is not that the display hurts my nose, I would really like to use it for longer time and I liked that I check the information proactively by turning my head, instead of something pops out on my monitor.* All participants gave positive comments about the gaze-contingent interaction and customization of widgets locations and notification levels. When asked if they were willing to use the interface on a daily basis if the form factor of the display was similar to a normal pair of eyeglasses, all participants gave a positive response.

**Perceived distractions:** All participants considered the Glanceable AR interface less distracting and disruptive to their primary tasks, compared to existing devices such as smartphones and personal computers. Participants commented: *I can easily check information without picking up and unlocking my phone, but they also stay out of the way when I don’t want them.* When asked if they thought the interface was distracting in some scenarios, three participants said they did not consider the interface distracting at all. They commented: *since I can customize their locations, I just put them slightly outside of my view, and when I find a widget gets into the way too often, I just mute it in the menu.* The other three participants commented that for the widgets that they do not want to use, even though
they are placed far away, they could still get in the view: *It would be great if I can make a widget disappear, or only see the ones that I need.* Two participants mentioned the level of distraction varied depending on the context of use. They commented: *when I am relaxing, I don’t wanna read my emails, the notifications from it kinda distract me.* As for notifications, four participants thought having a blinking icon at the periphery was acceptable, while the other two thought that the blinking behavior drew too much attention.

**Virtual content addiction:** A number of respondents in the first study worried that when information becomes more easily accessible through AR glasses, they might become more addicted to digital information. As such, we asked all participants if they thought they checked information in the widgets more frequently than they would normally do with their phones. Five participants gave a positive response. One commented: *yeah I definitely checked the news and weather more than I would normally do*

However, two participants who were inexperienced with AR thought there was a novelty effect. They commented: *they look cool and I think that is why I looked more frequently.* We also asked if participants were worried about virtual content addiction in AR based on their experience of using this interface. Four participants gave negative responses based on the current interface, but all of them said that addiction would be a concern if more widgets were supported. One participant said: *I cannot browse internet, watch videos, or check social media, but if I can, I could definitely see myself addicted with the glasses.*

**Follow mode and FG:** All participants were positive about follow mode. Participants commented: *I like having the widget around me but not on my face, and it allows me to keep getting information when I am moving.* Three participants mentioned that with the HG interface, widgets sometimes blocked the real-world environment. Three participants
reported that they used FG frequently, and all of them considered it more convenient and less obtrusive than the HG technique. One participant commented: *it feels magical that I can see through the dots without making the widget appear, and it is definitely less distracting than the other technique (HG)*. The other three participants reported that they did not use FG frequently due to inconsistent eye tracking.

**Favored scenarios:** All participants said that Glanceable AR provided benefits when working in front of the computer. Three participants reported that the interface allows them to decide quickly what to do without context-switching. One participant said: *when I get a email, I can quickly glance at the subject and see if it is important. If it is, I quickly go to my computer and reply to the email. In the most cases that it’s not, I just ignore it and go back to work, and it takes less than two seconds for me to decide.* Cooking was mentioned by four participants as a useful scenario. One said: *in one session when I was cooking my hands are dirty, and I do not have my phone with me, then an important email comes in, I might miss that email without the display.* As for scenarios they did not try but think the interface could be useful, three of them mentioned having a notes widget during a presentation. Two participants mentioned walking outside with the email widget to avoid missing important emails.

### 6.3.7 Summary of Findings

We summarize our findings as follows:

- All participants considered our Glanceable AR prototype easier and less distracting to acquire everyday information than existing devices.

- All participants mentioned that they would like to use Glanceable AR everyday if the
form factor of the display is ideal.

- Participants spent a short period of time paying attention to the AR information. They also developed their own strategies of using our systems in authentic everyday contexts.

6.3.8 Discussion

In this section, we will discuss our results based on the RQs.

**RQ6.E.** In the second study, we were able to confirm the positive quality of user experience of Glanceable AR in authentic everyday scenarios. The interface received excellent SUS and UEQ ratings, and participants expressed willingness to use the interface every day (ignoring the form factor of the display). Given the limitations in form factor of the Magic Leap One headset, the fact that we still received 126.61 minutes more usage time than required from participants also demonstrated the positive usability of Glanceable AR. The major benefits pointed out by participants were that it was less distracting and disruptive to their focus in the real world, as a quick glance requires less context-switching than pulling out a separate device. Having information out of the way but always accessible made the Glanceable AR approach able to handle common unstructured events in everyday life, such as checking email, news, and weather without taking too much of users’ attention or cognitive effort. However, participants also wanted more interactions with the widgets beyond simple information access. Some of them wanted to reply to emails directly in the widgets or create a calendar event by voice. Incorporating a higher level of interaction in Glanceable AR could contribute to a more useful experience, but might also lessen some of the benefits described above. The always-on property of the interface also received some criticisms. It could still be
distracting in scenarios when users want to be relaxed and isolated from digital information. Having an option to turn off the display completely might be helpful.

For follow mode use cases, occlusion problems in HG interface were mentioned by participants, which is not unexpected, and that is why we developed the FG technique. Although not every participants were able to experience it due to eye-tracking limitations, the fact that for the three participants that the tracking worked well, they all preferred FG over HG. This has motivated us to believe in the potential of applying gaze depth to avoid occlusions of virtual information on top of the real world. For wish-to-have functions, five participants mentioned they would like social media widgets to receive notifications from social platforms or YouTube widget and watch videos. However, as mentioned by participants, incorporating these widgets are likely to cause addiction to virtual information, especially given the ease of access. Careful design considerations have to be made to balance ease of access to information and potential addictions. Three participants mentioned that they would want to be able to send an email or create a calendar event directly in AR. One participant mentioned he would want to use hand gestures instead of the controller to place the widgets.

**RQ6.F & RQ6.G** We found that users performed quick glances more often than long glances. This indicates that participants were able to understand information in the widget and make decisions quickly within a short duration. The scenarios that users perceive as the most beneficial for Glanceable AR were very similar to what we showed in the video prototype in the first study. When viewing a primary physical display, participants demonstrated patterns of placing widgets around the physical monitor or clustering them by usage frequency. We found it very interesting that some participants developed the strategy of enlarging and placing the widgets all around the house to bind the widgets with the context of rooms. Combining Glanceable AR with context-sensitivity, it could be possible to enable
a more adaptive and unobtrusive information display.

Overall, our findings demonstrate that the Glanceable AR approach has great potential for everyday information access tasks. Our interface provides design inspiration for future implementations of Glanceable AR interfaces. Based on the feedback from participants, we would suggest incorporating gaze-contingent interaction features to maximize unobtrusiveness; integrating a follow feature for on-the-go access to widgets; empowering customization of locations, scales, and notification levels of widgets; and allowing users to dim the display completely when they do not want to be disturbed. We have confidence in the validity of our findings since they were obtained in unsupervised real-world settings in authentic scenarios.

6.3.9 Conclusion

In this research, we explored Glanceable AR as an approach for pervasive, general-purpose information display with future AR HWDs. In our first study, we demonstrated potential everyday scenarios of using Glanceable AR in a video prototype and obtained feedback from sixty-three participants in a survey. In the second study, we implemented a working application and studied real-world authentic use by six participants. To our knowledge, this was the first in-the-wild AR study that explored AR HWDs for general-purpose everyday use cases. Our results shed light on the design of future general-purpose interfaces for AR HWDs that are not distracting and easy to access at the same time. We found initial evidences for the potential of the Glanceable AR approach to support the everyday information access needs of average users once always-on AR glasses are widely available. However, as indicated in the future work section, more research has to be done to reveal longitudinal benefits of Glanceable AR systems, with more participants longer term of uses, and richer features to be versatile in all sorts of everyday contexts.
6.4 Capstone Implementation II: Evaluations of a feature-rich system

In the first capstone implementation, we were able to confirm the positive responses from participants about the initial Glanceable AR prototype in authentic everyday uses. Our prototype incorporates the interfaces we developed in chapter 3 and chapter 4 to resolve the potential distraction and occlusion issues of AR content on everyday users, while allowing users to customize their needs in terms of content, fixation, and notification. Participants considered it less distracting to use than existing mobile devices and would like to use it everyday if the form factor of the AR HWD is ideal. However, they also considered the system limited in terms of the features and applications. In this section, we aim at further improving the Glanceable AR system through blending in more of our research products in chapter 4 and chapter 5, and implementing a final proof-of-concept Glanceable AR prototype with enhanced features and interactivity. In the end, we conducted an expert evaluation of our prototype from twenty sessions of use over the course of three to five days. We detail the usage patterns, user behavior and experience, as well as lessons learned for future deployment of interactive everyday AR systems.

6.4.1 Capstone II Design & Implementations

Supporting more interactive features

We worked on including more features in each of our existing applications. We started by brainstorming and filtering through a list of interactive features. We narrowed down to a subset of promising features, focusing on quick and simple interactions that are frequently needed right after an information acquisition behavior (e.g., throw an email to the trash
6.4. CAPSTONE IMPLEMENTATION II: EVALUATIONS OF A FEATURE-RICH SYSTEM

Figure 6.8: (a-c) The new designs of the AR applications with added interactivity; (d-e) the new menu design with hand trigger and gaze interactions (i.e., Dwell, Blink) for selections.

folder after receiving it; check attendee and locations of a calendar event). As shown in Figure 6.8 (a-c) and Table 6.2, in the new version of the prototype, we included browsing features in the Gmail, Calendar and News apps (see Figure 6.9 (a)). The users were able to navigate through multiple emails, calendar events, and news pages through gaze-based dwell or blink interactions. In the Gmail app, users were able to trash an email or star an email after reading it (see Figure 6.9 (b)). We also kept the interactivity of the Clock and Task apps from our first prototype. As such, five out of the seven applications allow functions beyond viewing and consuming information.

Lowering the cost of manual interactions

Taking the learning from chapter 5, we attempted to lower the cost of manual interactions. In our first prototype, manual interactions such as triggering the follow feature, switching between interaction methods, muting certain applications, and placing applications were
Table 6.2: Interactive features added to selected applications.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Interactive Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gmail</td>
<td>Navigate/Expand/Trash/Star Emails</td>
</tr>
<tr>
<td>Calendar</td>
<td>Navigate/Expand Events</td>
</tr>
<tr>
<td>News</td>
<td>Refresh/Expand News Articles</td>
</tr>
<tr>
<td>Clock</td>
<td>Set Timer/Alarms</td>
</tr>
<tr>
<td>Task</td>
<td>Check/Uncheck Task Items</td>
</tr>
<tr>
<td>Weather</td>
<td>-</td>
</tr>
<tr>
<td>Activity</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 6.9: (a) The new design structure for the Gmail, Calendar, and News applications. Two arrows appear on top of the digital apps for browsing through multiple pages. Gazing at an interactive item (e.g., an email subject) would enlarge it, making it easier to glance at. (b) After reading the email, the user could glance at either the Trash or the Star buttons to perform the actions correspondingly; (c) Each app has a built-in menu with three features: follow (left), mute (middle) and close (bottom).

achieved through a hand-held controller. Although in our initial ideation of Glanceable AR, controllers and hands were considered unfavorable to use in everyday scenarios, they have been essential for 3D interaction and manipulation tasks in immersive environments. In this version of the prototype, we attempted to minimize the use of the hand and controller, with the controller only needed for initial placement of AR applications in the physical space (future systems would likely use bare-hand interaction for this purpose, but the ML1 does not fully support such interactions). The placement is automatically restored if the space is recognized, so it only needed to be done once for uses in the same space. Hands were used for rapidly summoning a menu, which was demonstrated to be advantageous in chapter 5.
as a quick way to open/close AR applications. In this prototype, after the user raises their hand and faces its palm towards the headset camera, a menu appears alongside the palm (see Figure 6.8 (d-e)). In this menu, users could toggle the visibility of each AR application, switch between HG and FG-based interaction methods, and enable/disable blink-based confirmation (see the following section) through gaze-based interactions. Functions that were universally applicable for each application, such as toggling follow mode, mute, and close, were moved to a universal menu that appears after the user gazes at the icon of each application (see Figure 6.9 (c)). As such, the uses of controller and hands were further minimized.

**Blinking as a confirmatory input**

In chapter 4, our results strongly favored Blink as an interaction method for rapid activation of AR content with high social acceptance. As a result, in this prototype, we implemented Blink as an alternative to FG to activate applications (see Figure 6.10). To reduce visual complexity, we replaced the visual indicators for visual feedback of blink detections with wiggling animations.

Furthermore, we attempted to integrate Blink, not only as a method of activation, but also as a confirmatory input. After blinking mode is enabled, the user can utilize blink as a replacement for dwell to prevent “Midas Touch” [82]. For example, instead of dwelling on an email title for confirmation of intent on viewing the body of the email, users could blink their eyes twice while looking towards the email subject. As such, users could take their time reading the subject of the email, not worrying about it being opened unintentionally after the dwell timer runs out. Through this change, we wanted to study the pros and cons of using Blink more universally as a confirmation interaction.
Figure 6.10: (a) The new design of the Follow feature. An icon replacing the dot to enhance awareness of the minimized applications; (b) A wiggling animation was used to indicate successful detection of user blinks; (c) After two consecutive blinks, the app was activated.

### 6.4.2 Research Questions

In implementing and evaluating a second capstone AR prototype, we aimed to answer the following research questions:

- **RQ4.H.** What features and design elements make the AR system practical for all-day everyday uses?

- **RQ4.I.** How do users perceive using a feature-rich Glanceable AR system for accessing, managing and interacting with AR apps?

### 6.4.3 Participants & Procedure

The study procedure was similar to the first capstone study. We recruited three participants. The study was divided into five phases. In the first phase, participants were asked to complete a background questionnaire and grant the AR application to access their personal account. In the second phase, an onsite tutorial session was provided to participants to walk through the hardware, calibration processes, the mobile app, and the Glanceable AR interface. In the next phase, participants were asked to use the AR HWD freely for at least six sessions
of at least 30 minutes, and to fill out a diary survey immediately after completing each session. Similar to the first capstone study, we chose 30 minutes empirically because it was a suitable time to wear the Magic Leap 1 AR headset without feeling uncomfortable or visual fatigue, so participants could ignore hardware constraints as much as possible while offering feedback. In the diary, we asked the participants about the time period, scenarios of use, layouts of the AR apps, which interactions they tried, and the overall perceived user experience in that session. In the fourth phase, after all sessions were finished, participants were asked to complete some post-study surveys (SUS and full version UEQ). In the final phase, we conducted a half-an-hour final interview with participants to ask in detail about their experiences of using the AR apps.

6.4.4 Measures

The quantitative measures we collected include the time participants spent (1) using the Glanceable AR prototype; (2) looking at and interacting with at each of the AR applications; and (3) using each given interaction method; as well as (4) the Likert scale data from the SUS and UEQ questionnaires. The qualitative measures we collected include diary entries and transcripts of the interviews. We also collected behavioral data include how the apps were arranged around the users.

6.4.5 Participants

We recruited three participants (1F/2M) from the local university (Mean age = 34). All participants self-identified as experts in AR and used Google services for daily work. All participants used smartphone, smartwatch, PC, and virtual assistant regularly.
6.4.6 Results

Usage sessions We collected 20 sessions of use from all participants, covering 624.52 minutes of use (10.41 hours). Participants used the prototype under a variety of scenarios, including working in a laboratory/office (13 sessions), in a class/meeting (5 sessions), having conversations with someone (5 sessions), walking around (2 sessions).

Usability Score The new Glanceable AR system obtained an average SUS score of 75 (SD = 2.5) from the three AR experts. Figure 6.11 illustrates the average UEQ scores compared to the benchmark scores. Other than perspicuity (M = 1.92 - Good), all the other categories received excellent ratings in the range of the 10% best results [139]. At the beginning of each diary entry, we asked participants to rate their level of user experience. All ratings were above average (≥ 5 in a 1-10 scale) with a mean rating of 7.58.

Usage behaviors On average in each 30-min session, participants used the AR widgets for 204.11 seconds (around 3 minutes and 24.11 seconds) (see Figure 6.12). Participants spent 66.50 seconds viewing/interacting with the Gmail app, followed by Google Calendar (42.35 seconds), Clock (21.78 seconds), News (20.79 seconds), Weather (19.50 seconds), Task
6.4. **Capstone Implementation II: Evaluations of a Feature-Rich System**

Figure 6.12: The average time participants spent on viewing/interacting with each of the AR apps.

(18.59 seconds), and Fitness (14.61 seconds). Participants spent 22.02 seconds opening/closing/muting apps. On average in each session, participants initiated and completed around 20 interactions with the AR applications.

The strategies participants used for arranging the AR applications included: (1) placing AR applications around the edge of a physical monitor; (2) arranging AR applications by categories; (3) placing AR applications around a physical desk; and (4) placing the relevant apps in easy to access locations and either closing the irrelevant apps or placing them further away in locations that are not visible (P1 & P2).

**Stationary Uses - Blink vs. Dwell** When user used the prototype in a stationary setting, they can choose either Blink or Dwell as the confirmatory input to trigger certain interactions (i.e., delete email, set a timer). Our results show that in the 30171.30 seconds (80.52%, 502.855 minutes) of using the prototype without moving around, participants used Blink for 16349.1 seconds (272.49 minutes, 54.19%), which is more than Dwell (230.37 minutes, 45.81%) (see Figure 6.14 (a)). Participants commented: “If I am trying to read an
email, I don’t want it selecting that email, like I wish that there is some way to let the system know that ... I want to interact with it. For that reason, with Blink, I can look at information without activating it, and if I want to interact with it, I just blink my eye twice (P1);” and “It works every time ... it was cool, and it was easier for me to access my apps [with Blink] (P2).”

**Mobile Uses - HG vs. FG vs. Blink** Throughout the 20 sessions, participants spent 7299.6 seconds (121.66 minutes, 19.48%) using the follow mode with at least one application following them (15.85% with one app, 19.60% with two apps, 26.65% with three apps, and 37.90% with four apps) (see Figure 6.13 (a)). Among the 121.66 minutes of follow mode uses, 3903.6 seconds (65.6 minutes, 53.48%) spent on using the HG interface, 710.7 seconds (11.85 minutes, 9.74%) spent on using the FG interface as the activation technique, and 2685.30 seconds (44.76 minutes, 36.79%) spent using Blink as the activation technique (see Figure 6.14 (b)).

**Open/Closing/Muting Apps** A difference between this prototype and its previous version is that we gave participants the capability to close the apps entirely at will. Figure 6.13 (b) shows the percentage of time with different amount of apps were opened. Our results show that 67.66% of the time participants did not close any apps. Then, 6.9% / 6.51% / 18.74% of the time participants closed 1 / 2 / 3 apps accordingly. The time that participants closed more than three apps were below 0.2%.

When it comes to muting apps, as shown in Figure 6.13 (c), 87.09% of the time participants did not mute any apps, and 12.88% of the time participants muted 1 app. The total percentage of time that participants muted more than 1 app is below 0.03%.
6.4. **Capstone Implementation II: Evaluations of a feature-rich system**

Figure 6.13: (a) The percentage of time in which users had 1/2/3/4 apps following them; (b) The percentage of time in which the users had 7/6/5/4 apps opened (the percentage of time that the users closed more than 3 apps is below 0.2%); (c) The percentage of time in which the users enabled 6/7 apps notification (the percentage of time that the users muted more than 1 app is below 0.03%).

**Distractions** When asked about the general user experience of the Glanceable AR prototype, all participants replied positively. Participants liked it because of its non-distracting and focusing characteristics: “I did not find the apps at all distracting (P1);” “Unlike when I normally work on my computer, I could .. have just the one main thing that I was working on, and then I can keep my calendar, email, and to-do list, and the clock using the AR apps (P2);” and “It helps me to focus on my main screen, and also being able to access all the apps at the same time without having to leave my main screen out of my field of vision (P3).”

**Interactive Features** When asked about their perceptions of the interactive features of the applications, all three participants considered them easy to use, helpful, and non-distracting. Participants commented: “When I get an email, I can quickly glance and delete it if it is not important (P1);” “I used the trash feature a few times, which was nice because I did not have to open up my email client to do that (P2);” “I really like the fact that we can interact with those apps ... the possibility that we can interact with the apps and like see
things and trash emails and everything ... make it easier in such a way that I do not have to
go to my computer. [I can] interact with the things I see just in AR (P3).” Participants also
wished for more features and app support: “If we can have some pre-defined answers like
Gmail has, like thank you, sounds great, if I can blink on them to send, it would be super
helpful (P3);” “I wish I can have a music control app to switch songs.” When being asked
whether they would like to use such a system daily if the form factor of the AR display were
ideal, all three participants gave a positive response.

6.4.7 Summary of Findings

We summarize our findings as follows:

- All participants favored the interactive features in the new Glanceable AR prototype.

- All participants mentioned that they would like to use Glanceable AR everyday if the
  form factor of the display is ideal.

- Participants liked using the Glanceable AR prototype when having social encounters
with others.

- Participants considered Blink to be more useful than Dwell as a way to activate or trigger interactions.

### 6.4.8 Discussion

In this study, we implemented a second Capstone prototype of our Glanceable AR system, in which we improved the level of interactivity and applied the knowledge gained throughout this dissertation. Our results demonstrated, again, the positive impact of using AR in everyday lives while being supportive, non-distracting, and easy to use even with the limitations of current hardware. Our prototype received high usability ratings from our expert users. Participants considered the approach novel, attractive, efficient, dependable, stimulative, and helpful. The most successful scenario of using the Glanceable AR system, from both studies, is when users had a primary task or objects that needed attention, in which they would customize the locations of the AR apps to be around the edge of the physical object. For example, when users were working in front of a desktop, having a class, or talking it someone. In such cases, they could quickly grab information with a glance whenever needed and quickly go back to their primary tasks with little to no context switching.

The added interactivity was deemed useful. They increased the time participants used the AR apps as compared to our Capstone I study, but did not hinder the unobtrusiveness of the AR apps. In contrast, participants complimented that the added interactivity allowed them to further perform quick actions without pulling a secondary device or web application, which kept their workspace clean and focused. Participants also mentioned that they would like more applications and features. Although it is possible that with richer interactions, we could better support participants’ tasks. However, this may come with more distractions,
similar to what our mobile phones offer nowadays. Participants may also be overloaded with interaction possibilities, as well as the perception and cognition costs to make an interaction decision. Design choices need to be carefully considered in order to strike a balance between the level of interactivity and the potential distraction and obtrusiveness levels.

Since our first Capstone evaluation was conducted during the global pandemic, all the sessions we collected were single-user use case without social encounters. In this study, we were able to collect diary entries detailing scenarios when participants were having face-to-face conversations with others wearing the AR HWDs with the AR applications following them around. Participants commented that they felt the Glanceable AR system was socially-friendly: “Normally, it would be considered rude to check your phone while talking with someone. However, with this system, I was able to check the email as it was coming in without seeming rude (P1);” “While having conversations with people, it’s way nicer to use the Glanceable apps for quick checks (of email, time, calendar) compared to taking out my phone or even looking at my watch. If I had used my phone/watch the same amount that I used the Glanceable apps, people would have been annoyed with me, or thought I wasn’t paying attention to them (P2).” This further proved the advantages of unobtrusive interfaces in AR HWDs. The unnoticeability and easy-accessibility made it viable to support social conversations without being as in the way as mobile phones and smartwatch. However, some potential issues were also mentioned: “I probably checked my email too much during the conversations, and so I wasn’t as completely present as I should have been (P2);” and “since the therapist [conversation partner] cannot see my eyes ... it might be difficult for her to assess my reactions and emotions. Also, there’s always a possibility that I get distracted by a new notification (P3).” Due to hardware limitation, users found it hard to keep eye-contact with the conversation partner. On the downside of high accessibility, users may feel tempted to be drawn away by the presence of the AR apps around them. A solution to this,
as mentioned by one participants (P3), is to enable the users to author profile modes, such as work mode, productivity mode, meeting mode, to avoid irrelevant apps from appearing and consuming users’ attention unnecessarily.

In general, participants used Blink more than Dwell in stationary uses, and used it more than FG for mobile cases. The primary reason of favoring Blink, as mentioned earlier, was due to its controllability and confirmatory nature: “there are times where I just want to look at the displayed information without trying to activate the various components. I wish there was a discreet way that I could tell the system that I am ready to interact via eye gaze. The blink condition sort of gets close to this (P1).” However, participants also experienced difficulties using Blink due to tracking issues: “the current implementation of blink is not highly reliable for me, so I often had to try multiple times to activate something (P2).” During Blinking when users’ eyes were half-closed, the eye-tracking results jittered, which caused issues when the visual targets were small. Better tracking algorithms will be needed to stabilize the gaze cursor for Blink to be more robust and usable in everyday scenarios. In our studies, participants did not report fatigue while using the Blink technique. In previous research, blink interaction has been explored as a text-entry input for motor-impaired populations [141]. In their evaluations, participants reported slightly higher than average visual fatigue. This indicates that frequent prolonged use of blink interaction could cause eye strain. Future interface should be mindful of that by reducing the required blink interactions within a short duration.

### 6.4.9 Limitations and Future Work

There are several limitations of our work. In the survey study study, our sample was not gender-balanced, which could have affected the results to some degree. Second, participants’
responses after watching the video prototype might not reflect their opinions of actually using the interface. We addressed this limitation with the capstone evaluations. Third, the participants of the Capstone I study were all college students, which could contribute to working in front of a computer being the most frequent scenario of using the interface. Future work could be conducted to recruit more participants with diverse backgrounds to evaluate the approach. Fourth, several participants were not in the habit of using some of the widgets in their daily lives (e.g., three participants did not use the Google fit service), which made the interface less useful to them. Fifth, participants only used the interface indoors in home or office environments. Further research could explore the user experience of Glanceable AR outdoors. Seventh, interactions with the widgets including positioning and customization were achieved through a controller. The hand-held nature of the controller could make it challenging to be applied to everyday use cases. Future research could explore voice, hand gestures or gaze as potential input modalities. Finally, we did not collect any data about the environmental features that participants used the interface in. Future research could explore how environmental features (location, space, public/private) would influence patterns of using the interface. Last, both capstone implementations were conducted within the duration of three to five days. We used this duration due to physical and environmental constraints. We acknowledge that a longer term of use (e.g., 3 weeks or 3 months) may bring different insights to our results, allowing investigations on our system with more granularity, and revealing how user perceptions and behaviors evolve cross-sessions. Future study could evaluate everyday AR systems with longer-term of uses.
6.5 Conclusions

To conclude, in this chapter, we described our work on designing, developing, and evaluating proof-of-concept Glanceable AR systems. We surveyed public opinions of Glanceable AR by delivering it to a broad sample of population, and evaluated two Glanceable AR prototypes in authentic use cases. Our findings contribute to better understanding of the effectiveness of the proposed approach in this dissertation. We provide valuable insights and recommendations for designing unobtrusive and efficient information displays in AR for average users once everyday AR glasses become widely available. However, there are some limitation to this work. Due to time constraints, we were not able to integrate all the features that we researched (e.g., the Some transitional interface) and perform a large-scale longitudinal evaluation of the prototype with everyday users. Future research could look into evaluating Glanceable AR systems with spatial adaptivity, more users, longer-terms of uses, and with supports of world-anchored transitional interfaces for on-the-go uses.
Chapter 7

Conclusions and Future Work

7.1 Summary

AR HWDs have become more lightweight and powerful. Soon, they could be worn all-day, delivering virtual information continuously and pervasively to support people’s everyday tasks [56, 65]. However, very few studies have explored general-purpose uses of AR HWDs to support everyday information needs with users. With AR content appearing everywhere and becoming always-available, it could be distracting, obtrusive, and occluding real-world objects of importance. How can we design efficient, natural, non-distracting and relevant information displays? In this dissertation work, we focused on addressing these problems. We aimed to answer four research questions:

- How can we support non-distracting yet efficient and natural information acquisition with AR HWDs? In chapter 3, we described our initial concept of Glanceable AR, an interaction paradigm that allows unobtrusive acquisition of information through AR HWDs. We proposed three interfaces under the paradigm, and evaluated them through a user study covering two everyday scenarios in which different levels of attention is required on the digital content. Our results, although favoring a HUD-type display, offered insights about the benefits of placing digital content outside of the central FoV of the users. This led us to dive deeper into how to further minimize the obtrusiveness
of the virtual interfaces, and evaluate our solutions in more ecologically valid everyday settings.

- **How can we prevent AR content from being obtrusive and occluding the real-world objects of interest to the users?** From studies conducted in chapter 3 and chapter 6, we found evidence favoring the Head-Glance interface. Participants considered it less distracting while still being convenient to access information. However, moving virtual information outside the central FoV cannot completely resolve the *occlusion* issues, given the use of glancing as a behavior for attention to not only the digital content, but also the physical environment. After a preliminary study that confirmed this issue, in chapter 4, we studied interactions to further improve the unobtrusiveness of digital information. Leveraging gaze, hand, and head-based input modalities, we developed strategies for users to disambiguate their intent between the digital and physical layers. Through a study simulating information acquisition tasks in both the real world and virtual content while seated or walking, our results provide empirical insights on when and how to deploy user-initiated interactions for quick access to either the digital or the real world without compromising the accessibility of the other.

- **How can we support relevant and scalable acquisition of AR content during spatial transitions?** In order to support people’s all-day tasks, AR interfaces need to support a variety of applications on-the-go, which have not yet been well-supported by existing AR devices. In current devices, AR applications default to be world-anchored after being opened, and need to be reopened when the users transition to a new space. In chapter 3 and chapter 4, we researched display-anchored and body-anchored ways of carrying AR information with the users for on-the-go access. A critical issue with such solutions is *scalability*. An increased number of applications could quickly occupy the users view and become overwhelming. In chapter 5, we dived deeper into the
“Fixed to Users” design principle of Glanceable AR, aiming to improve the mobility of AR applications when their quantity scales up. We investigated the trade-offs among three solutions: (1) None: an interface that prioritized visibility of the real-world; (2) All: an interface that prioritized accessibility of information; and (3) Some: an intelligent interface that aimed to improve accessibility of information while minimizing the required manual input and the occlusions. Through two user studies, we modelled how the user’s performance correlates to the number of AR applications, and evaluated the three interfaces through a ecologically-valid role-playing study. Our results shed light on supporting scalable information access with AR applications when users move across multiple locations.

• What are the user perceptions and acceptance towards using pervasive and always-on AR systems in authentic everyday contexts? In the previous questions, we proposed interaction methods and proved the feasibility of them through empirical evaluations in the laboratory. However, there is still a lack of understanding on how everyday users would actually perceive using such systems in their daily lives. In the past decades, though there has been extensive research on developing AR/VR systems to support a wide variety of special-purpose tasks and use cases, few explorations have been done regarding how users would perceive using AR systems for general-purpose everyday task scenarios due to the form factor of hardware. With the recent advancements in technology, we have reached a point where such evaluations are feasible. In chapter 6, we conducted one large-scale survey study and two in-the-wild studies that reveals these use cases in an unsupervised manner to understand genuine user perceptions and requirements about using AR HWDs to access personal information in various everyday scenarios. Our results provided evidence of the potential benefits of our proposed approach, as well as insights about how to deploy AR systems that meant to
be pervasive, always-on, supportive, and unobtrusive to the user’s tasks.

Taking our findings across all four chapters covering eight user studies, we highlight the following core learnings:

- People hold positive views overall about the future of AR HWDs for everyday all-day uses, while being worried about digital information taking over their physical reality and becoming distracting, and invasive to their life.

- Future versatile general-purpose AR interfaces should always prioritize the real world while keeping virtual content easily available, accessible and supportive.

- Placing information in the periphery is an effective and favorable strategy to access or monitor secondary information for future AR systems.

- Gaze-based interactions, especially eye blinks, is considered favorable for activating minimized virtual content and interacting with it.

- Environment-based intelligent recommendations of AR apps could be an effective strategy to allow scalable access to AR applications during user transitions.

- Despite limitations of current hardware, our Glanceable AR approach, in which virtual information maintains unobtrusive and non-intrusive to the users physical tasks while being easily accessible, is considered useful and promising for everyday uses.

Overall, this dissertation contributes to wider adoptions of AR HWDs in the future by researching ways of displaying everyday information in AR appropriately, efficiently, supportively, and non-invasively in a variety of common everyday contexts. Through iteratively designing, implementing, and deploying interactions and rigid evaluations in user studies, we offered design recommendations for deploying AR HWDs in authentic everyday scenarios.
7.2 Synthesis & Reflections

In this section, I reflect on my five years of Ph.D. journey, what we have learned overall, what Glanceable AR represents, and how our findings hold as technology continuously advances.

7.2.1 Definition of Glanceable AR

The definition of Glanceable AR has gone through several iterations throughout this dissertation work. The initial definition, as demonstrated in subsection 3.2.2, was constrained in scope, and mainly took into consideration mobile on-the-go uses of AR information. Throughout the dissertation, our definition kept expanding and went to a broader scope. Taking into account limitations in current technology, our definition focuses less on interactions but rather the form of delivering information to the users (see subsection 1.3.1). In short, Glanceable AR refers to a information display strategy that puts AR information secondary, draws minimal attention from the users when unneeded, prioritizes the real world environment, leverages the periphery of either the display or the user’s field of view to show or store information, and is accessible no matter where the user is located. An AR information display would not be glanceable if it is the primary focus of the users, for example, using a virtual monitor for productivity scenarios or playing an immersive AR game. Similarly, a piece of AR content that was glanceable seconds ago may no longer be later after being prioritized. Glanceable AR refers to a state in which AR information remains secondary and auxiliary to the user’s primary task, either physical or virtual. In some ways, Glanceable AR resonates with a philosophy called digital minimalism, which advocates for living a focused life in a tech-saturated world by minimizing the engagement with the online worlds [120]. Similarly, we believe that by lowering the presence of unnecessary digital realities, users would be empowered to focus on the tasks that are of importance to them.
7.2.2 Potential impacts of perfect hardware

Due to limitations in current technology, AR HWDs are still heavy, cumbersome, has FoVs far lower than human visions. It is reasonable to ask, would our learnings still hold when AR HWDs reaches eyeglasses form factor? This is a tricky yet intriguing question. It is arguable that in some evaluations, the limitation of the hardware may have hindered some negative implications of the Glanceable AR approach. For example, the small FoV of the Magic Leap 1 and HoloLens 2 headsets may have made virtual content in the periphery more invisible. If the device has FoV close to human eyes, the users may notice the AR content in the periphery more easily and be more tempted to look at it. When we reach that point, further research may be required to investigate where in the periphery should the AR content be located and how should we show them. However in general, we believe that we have designed study scenarios with high ecological-validity, assessed our interfaces in-the-wild with actual users, and managed to minimize the negative implications of current hardware. The fact that we were able to obtain positive feedback from participants about our systems even with current limitations also demonstrated the promising potential of our approaches. We believe that most of our lessons learned would last long, and re-evaluations of the Glanceable AR approaches would be an important next step once we reach to that point.

7.2.3 Impact of the pandemic on this dissertation

Having COVID-19 as part of my Ph.D. journey was a unique experience. The pandemic made life harder for many HCI researchers, especially in the AR/VR community, where in-person evaluations of prototypes or proposed methodologies were often mandatory. When the pandemic hit In 2019, I was just about to run an in-person study evaluating Glanceable AR in
a more demanding simulated task. If the Ph.D. is like sailing on the ocean, then the pandemic was like a hurricane, blocking the planned sail path and forcing a detour in my Ph.D. research. However, I was lucky that the detoured path shined with possibilities. We had to suspend our planned study, and worked on conducting remote evaluations of the Glanceable AR approach. This gave birth to research in the entire chapter 6. Being quarantined at home allowed us to conduct adequate self-evaluations and iterations of the Capstone Glanceable AR prototype, and forced us to conduct unsupervised in-the-wild evaluations of our systems. The evidence we gained in favor of our system encouraged me to continue my research on Glanceable AR and eventually finish this dissertation work. Due to the pandemic, I had to change my path, missed countless opportunities to meet fellows and colleagues and present my research at conferences in person, while we also set paths on research possibilities that we may not have considered otherwise.

7.3 Future Work

I would like to point towards future research directions that could be complementary to this dissertation work.

7.3.1 AR information displays

This dissertation work sets paths to unobtrusive AR information displays, though with its limitations. First, although we explored head, body, world, and surround-fixed layouts in our work, we acknowledge that there are more nuances in layout strategies that are under-explored. For example, future research could investigate ways of automatically transitioning between body-fixed and world-fixed layouts based on environment and user activities, as
presented by Ens et al [52]. The AR app could also embody multiple layout strategies. For example, they could be world-fixed, but also billboard towards the user’s head orientation for easy readability [56]. Second, this dissertation explores mostly AR overlays, in which AR content is displayed as 2D windows, registered to but somewhat isolated from the real-world background. We believe that this is a paradigm that people are more familiar with interacting with 2D screens such as laptops and phones, and a more feasible way of showing information with current technologies. However, one may argue that this does not leverage the benefits of AR to the fullest, since the strength of AR devices is its ability to display information in more situated and integrated forms with high relevance to the real-world objects and environments. For example, a recipe app may directly outline ingredients instead of showing the list in a panel, a navigation app may augment the real world with glowing pathways and x-ray vision for intuitive guides [79, 80, 94, 129]. However, these strategies push for more advanced tracking, sensing, IoT, and computer vision support. Research is also needed to further identify their perceptual and cognitive benefits as compared to traditional 2D information display approach in different everyday scenarios.

Furthermore, in our study, although we attempted to study Glanceable AR with information presented at the periphery, due to limitations of the FoV in current AR HWDs, the user’s peripheral vision was not fully leveraged for efficient peripheral awareness. This issue could be potentially resolved by using projection-based AR displays or simulated-AR with VR headsets, which though would made the system impractical to use in actual everyday use cases due to the extra weight, sensors, and hardware. We call for future research studies that fill this gap by studying placements of AR information in the user’s periphery.
7.3.2 Cross-device interactions

AR HWDs offer great potential to fulfill everyday information needs, as demonstrated in this dissertation work. However, due to limitations in hardware, battery, and optics, they are not fully ready as a versatile everyday device in their current state. Current AR devices suffer from small FoV, cumbersome mid-air input, lack of tactile feedback and situated interactions. With these constraints, AR HWDs may not be able to fully replace existing mobile devices within a short time frame. Instead they may complement the existing mobile devices ecosystem. For example, an everyday AR user could use AR HWDs for quick consumption of information, and use mobile phones for precise input such as text-entry, pointing, and selecting targets [64, 171]. Similarly, a smartwatch could be used for sensing finger tapping, wrist and hand movements given its rich IMU sensors for free-hand gestures and pointing tasks [112, 147, 168]. A laptop would be suitable for more dedicated tasks, which AR devices and phones are less capable of due to the lack of software infrastructure. However, the user could easily leverage AR devices for concurrent viewing and editing in the physical world. One may argue that as AR devices keep advancing themselves, they may become the “ultimate display” that eliminate completely the need for other physical devices, thus research in cross-device interactions may be less applicable as time progresses. However, such research may be necessary to accelerate the proliferation and adoption rate of AR devices when they are not yet perfect, and leave insights for interaction metaphors and affordances that people favor when AR HWDs become the sole needed display that fulfill people’s computing needs anywhere and anytime.
7.3. Future Work

7.3.3 Social issues with AR HWDs

A critical yet unaddressed challenge with AR HWDs is social acceptance. In this dissertation, even though we touched upon social and conversational use cases in chapter 5, our evaluations primarily occurred in a single-user use case, in which only one user had access to the AR display. However, in the envisioned future of pervasive AR, everyone would have access to a pair of AR glasses. This brings tremendous opportunity for social interactions such as content sharing. However, certain issues may arise due to the lack of social cues. If taking interactions with mobile phones for examples, holding the phone, looking at it, and using the touch screen are obvious and clearly visible to co-present others indicating that the user’s attention is partially drawn to the digital realm. However, for all-day head-worn AR displays, the boundary between physical and digital realities may dwindle. Each user may have their own pair of AR glasses displaying dedicated virtual information that is invisible to others. Head and gaze-based interactions have been considered as promising ways to interact with future AR displays. Such interactions are subtle because we move our head and eyes involuntarily in everyday activities all the time. For example, imagine an AR glasses wearer looking in your direction through the lenses, they could be reading an email in AR, they could be zoning out, they could be attempting to get your attention and say Hi, or they could be in a virtual meeting and do not want to be disturbed. Such ambiguity brings challenges to interpersonal interactions in social situations. As such, future investigations are demanded in order to develop interactions and interfaces that are socially-supportive and privacy-preserving.

Another interesting direction is how to design such interfaces that better support collaborations. In collaboration activities, the users often need both private and shared spaces. They may also need information that involves only a fraction of each other’s AR UIs. For example, users may need different visibility levels (e.g., whether their collaborators are looking at the
Figure 7.1: If user A is reading a piece of virtual content, Glanceable AR interfaces could vary in translucence levels that visualize user A’s interactions to user B to achieve more efficient and comfortable collaborations (the blue diamond indicates that the user is currently looking at a piece of digital information rather than the physical world).

real world or virtual information, where is the virtual content, the high-level theme of the virtual content, and the detailed information in the virtual content) to collaborate more fluidly with each other (see Figure 7.1). How to design Glanceable AR systems that allows users to control the translucency of their content/interactions to better support collaborations [53], would be a critical next step.

### 7.3.4 Context-aware AR systems

This dissertation investigated interface solutions that were mostly user-initiated. Such strategies ensure high controllability and stability, which are critical to ensure a satisfying user experience in dynamic everyday scenarios. On the down side, user-initiated interactions require users to make the input manually. Recent research has highlighted system-initiated interactions that leverage the rich sensory capabilities of the AR devices to infer the user’s tasks and environment, or more broadly their “context”, in order to provide the right service,
at the right time, in the right place, right when the users need them \([98, 101, 169]\). In chapter 5, our result shed light on such possibility by highlighting the benefits of a location-based automated interface. However, AR devices grant access to user data beyond physical locations. They are worn on the head and share the same perspective with the users for viewing and interacting with the real world. Recent AR/VR/MR systems come with internal sensors for eye-tracking, hand-tracking, and face-tracking. It remains an open question how to leverage the rich data gathered from the users to create a more satisfying user experience. We would like to call for future research into studying ways of leveraging the data gathered by devices for context-aware and adaptive glanceable information displays, enhancing further the unobtrusiveness and relevance of AR information. For example, imagine if the AR system could detect that a person is behind an AR widget and automatically minimize or move it \([43, 80]\); or that it could automatically adapt the layout of the applications depending on the user’s mental load, task, and environment \([30, 54, 128]\); or that it could adapt the visual representations of a given AR app (e.g., shape, size, 2D vs. 3D, overlay vs. integrated) depending on use scenarios for lower cognition and perception costs. Recent research has investigated these possibilities by optimizing a cost-benefit model through a limited set of parameters \([30, 98]\). However, events in the real-world environment are far more complicated, with unexpected beginning/ends, interruptions, and dynamic scenes. This creates unique challenges for wide applications of automated interfaces to everyday scenarios. An everyday task scenario may involve a multiplicity of long-term tasks which may be changing in relative level of performance. Further research is needed on how AR systems may detect the attention of the users, learn from the user’s historical interactions, determine the priority of users’ tasks, and deliver the right information to them in more seamless ways.

The two critical challenges that remain unresolved with adaptive UIs, as highlighted in prior work, are consistency and controllability \([57, 136]\). A highly adaptive UI disrupts spatial
consistency and the user’s mental model of the UI’s locations and functionalities. It also diminishes the sense of agency, which has been proven to be critical for ensuring a satisfying user experience in automated interfaces. As such, extensive research is needed in human-AI interactions for AR/VR systems to enable a responsible, trust-worthy, non-intimidating, and co-existing future [8, 34]. A potential best-of-both-worlds solution to this may be mixed-initiative, or adaptable interfaces, in which the system suggests adaptation methods that maximize computational benefits for the users, but it would be up to the user whether or not to adopt it [57, 101]. Recent research has also found that user’s personality traits would interfere with their perceptions of interacting with conversational AI recommenders [26], which further solidified that there is hardly one-size-fits-all AI solution or algorithm for everyday users. Future context-aware AR system should learn from its users, involve them in the loop, and continuously advance itself through learning the users’ reactions and behaviors.
Bibliography


[29] Cen Chen, Chilin Fu, Xu Hu, Xiaolu Zhang, Jun Zhou, Xiaolong Li, and Forrest Sheng Bao. Reinforcement learning for user intent prediction in customer service


[34] Nazli Cila. Designing human-agent collaborations: Commitment, responsiveness, and


[67] Jonna Häkkilä, Ashley Colley, and Juho Rantakari. Exploring mixed reality window concept for car passengers. In *Adjunct Proceedings of the 6th International Conference*


[128] Xun Qian, Fengming He, Xiyun Hu, Tianyi Wang, Ananya Ipsita, and Karthik Ramani. Scalar: Authoring semantically adaptive augmented reality experiences in


1007/978-3-642-24198-7_1. URL https://doi.org/10.1007/978-3-642-24198-7_1.


[169] Xuhai Xu, Anna Yu, Tanya R. Jonker, Kashyap Todi, Feiyu Lu, Xun Qian, João Marcelo Evangelista Belo, Tianyi Wang, Michelle Li, Aran Mun, Te-Yen Wu,


Appendices
Appendix A

Designing for Unobtrusive and Non-Distracting Information Access

A.1 Informed Consent Form
Title of research study: IRB #19-544: Exploring the trade-offs between Glanceable and Summonable Interactions for Secondary Information Access in Always-on AR Devices

Principal Investigator: Dr. Doug A. Bowman Contact: (540) 231-2058 or dbowman@vt.edu

Other study contact(s): Feiyu Lu Contact: (540) 257-4562 or feiyulu@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

This research project is intended to compare a variety of methods for interacting with augmented reality (AR) systems. This research will help us understand the best ways to interact with data and information in VR and AR and produce design guidelines for VR/AR user interface design.

Detailed Information: The following is more detailed information about this study in addition to the information listed above.

Who can I talk to?
If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at Room 160 (Sandbox), Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060. This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

How many people will be studied?
We plan to include about 48 people in this research study.

What should I know about being in a research study?
- Someone will explain this research study to you
- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind
- Your decision will not be held against you
- You can ask all the questions you want before you decide

What happens if I say yes, I want to be in this research?
You will become a participant in this research. You will be invited to try out the AR/VR system we developed, where your data will be collected to help us understand the best ways to interact with data and information in VR and AR. The study will take place in the Sandbox Studio (Room 160, Moss
Consent to Take Part in a Research Study

Arts Center, 190 Alumni Mall, Blacksburg, VA 24060), and will take approximately 90 minutes for each participant.

What are the experimental procedures?
Upon signing up, the participants will be asked if they have participated in previous related studies. When participants arrive, they will be greeted and asked to read and sign the informed consent form after their questions (if any) are answered. Then they will be asked to clarify they have normal vision, either uncorrected or corrected by contact lenses. Next, they will be provided with written or verbal instructions for the experiment and familiarized with the lab and the equipment they will be using. Participants will wear a Magic Leap One augmented reality (AR) headset. Using these devices, participants will then complete a series of interaction tasks in AR, using one or more 3D interaction techniques we designed. Tasks will involve physical movements including looking around the environment, walking to targets, pointing to objects, and manipulating virtual objects. After each block of tasks, participants will be interviewed by the investigator to gather qualitative feedback. Breaks will be given after each usability questionnaire. After all tasks are completed, participants will be interviewed about their experience. The experiment will be video and audio recorded for researchers’ further interpretations and analysis of user behavioral patterns.

What happens if I say yes, but I change my mind later?
You can leave the research at any time, for any reason, and it will not be held against you. All the data collected to the point of withdrawal will be properly discarded and destroyed immediately.

Is there any way being in this study could be bad for me? (Detailed Risks)
Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment, or contact the physical cables connecting the display to the computer. Each participant will be accompanied by an experimenter alongside as a safety guarantee.

What happens to the information collected for the research?
We will make every effort to limit the use and disclosure of your personal information, including research study data and video/audio recordings, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

Can I be removed from the research without my OK?
Participants will not be removed from the study without their consents.
What else do I need to know?
This research is being funded by Office of Naval Research via University of California, Santa Barbara.
We will not offer to share your individual test results with you. You may accept or decline these results.
Your information and samples (both identifiable and de-identified) might be used to create products or to deliver services, including some that may be sold and/or make money for others. If this happens, there are no plans to tell you, or to pay you, or to give any compensation to you or your family.
There are no significant benefits that individual subjects might experience from participating in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset.
Consent to Take Part in a Research Study

Signature Block for Capable Adult

Your signature documents your permission to take part in this research. We will provide you with a signed copy of this form for your records.

______________________________  _________________________
Signature of subject             Date

______________________________
Printed name of subject

______________________________  _________________________
Signature of person obtaining consent  Date

______________________________
Printed name of person obtaining consent
A.2 Questionnaires
Pre-study Questionnaires:

1. Gender
   a. Male
   b. Female
   c. Prefer not to tell
2. How are you feeling today?
   a. 1 (Negative) to 7 (Positive)
3. Do you wear contact lenses?
   a. Yes
   b. No
4. Please indicate your occupation (if student, please indicate undergraduate or graduate)
5. Major (if student)
6. How many times have you tried Augmented/Mixed Reality?
   a. Never
   b. Once or Twice
   c. 3–10 times
   d. More than 10 times
   e. Comments: ______
7. Which augmented/mixed reality devices have you tried?
   a. Mobile AR
   b. Microsoft HoloLens
   c. Magic Leap One
   d. Meta 2
   e. Windows Mixed Reality
   f. Comments: ______
8. How often do you use mobile phones to acquire general information everyday (e.g., weather, news etc)?
   a. Never
   b. Once or Twice
   c. 3–10 times
   d. More than 10 times
   e. Comments: ______
9. Do you believe that augmented reality will replace our mobile phones in the future?
   a. Yes
   b. No
   c. Comments: ______

SUS Questionnaire

1. I think that I would like to use this technique frequently.
   Strongly Disagree    Disagree    Neutral    Agree    Strongly Agree

2. I found the technique unnecessarily complex.
   Strongly Disagree    Disagree    Neutral    Agree    Strongly Agree

3. I thought the technique was easy to use.
4. I think that I would need the support of a technical person to be able to use this technique.

5. I thought there was too much inconsistency in this technique.

6. I would imagine that most people would learn to use this technique very quickly.

7. I found the technique very cumbersome to use.

8. I felt very confident using the technique.

9. I needed to learn a lot of things before I could get going with this technique.

10. I was physically comfortable using this technique.

**NASA TLX Workload Questionnaire**

1. **Mental Demand:**
   How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?

   1 (Very Low) – 7 (Very high)

2. **Physical Demand**
   How much physical activity was required? Was the task easy or demanding, slack or strenuous?

   1 (Very Low) – 7 (Very high)

3. **Temporal Demand**
   How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?
4. **Overall Performance**
   How successful were you in performing the task? How satisfied were you with your performance?
   
   1 (Perfect) – 7 (Failure)

5. **Effort**
   How hard did you have to work (mentally and physically) to accomplish your level of performance?
   
   1 (Very Low) – 7 (Very high)

6. **Frustration Level**
   How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?
   
   1 (Very Low) – 7 (Very high)

**Interview questions:**

1. Which part of the following interaction techniques that you tried today did you like or dislike? Why?
   - a. HUD
   - b. Head-Glance
   - c. Gaze-Summon

2. How did you feel that the interaction techniques affected the performance of your primary task? Why?

3. Could you please rate and rank the interaction techniques in your own preference?
   - a. HUD
   - b. Glanceable
   - c. Summonable
A.3 Semi-structured Interview Scriptss

A.4 IRB Approval Letter
MEMORANDUM

DATE:  July 3, 2019

TO:  Douglas Andrew Bowman, Feiyu Lu

FROM:  Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE:  Exploring the trade-offs between Glanceable and Summonable interactions for secondary information access in always-on AR devices

IRB NUMBER:  19-544

Effective July 2, 2019, the Virginia Tech Institution Review Board (IRB) approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at: https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Approved As:  Expedited, under 45 CFR 46.110 category(ies) 6,7
Protocol Approval Date:  July 2, 2019
Progress Review Date:  July 1, 2020

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.
<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td>06/06/2019</td>
<td>PET47EVN</td>
<td>University of California, Santa Barbara (Title: View management and user interface optimization for wide-area mobile augmented real)</td>
<td>Compared on 06/17/2019</td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.
A.5 Video Demonstrations

Video demonstrations of this research is available both online here and in the supplementary material files that are attached to this dissertation. The presentation recording of this paper is available here.
Appendix B

Resolving Occlusion Issues of AR Content

B.1 Preliminary Study

B.1.1 Informed Consent Form
Title of research study:_IRB #21-090: Identifying and resolving occlusion issues for Glanceable AR Interfaces

Principal Investigator: Dr. Doug A. Bowman  Contact: (540) 231-2058 or dbowman@vt.edu
Other study contact(s): Feiyu Lu  Contact: (540) 257-4562 or feiyulu@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.
We invite you to take part in a research study about Augmented Reality (AR) user interfaces. The research will take around 20 minutes to complete.

What should I know about being in a research study?
- Someone will explain this research study to you
- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind
- Your decision will not be held against you
- You can ask all the questions you want before you decide.

What should I know about this research study? This research project is intended to compare a variety of methods for interacting with augmented reality (AR) systems. This research will help us understand the best ways to interact with data and information in AR and produce design guidelines for AR user interface design. You will be asked to wear a Magic Leap One AR headset and interact with several pieces of virtual content displayed in the headset while sitting or walking inside a room. Afterwards, you will be asked to fill out questionnaires about your experience (More detailed information can be found under “What are the experimental procedures?”). The research will take around 20 minutes to complete. There is risks that you might feel dizziness or nausea (More detailed information can be found under “Is there any way being in this study could be bad for me? (Detailed Risks)”). There are no significant benefits that individual subjects might experience from participating in this research. We cannot promise any benefits to others from your taking part in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset. Potential benefits to others also include the AR industry gaining a better understanding of occlusion issues in everyday AR uses. If you are a student at Virginia Tech, the
Consent to Take Part in a Research Study

decision whether to participate or not will have no effect on your academic performance or relationship with Virginia Tech.

**Detailed Information:** The following is more detailed information about this study in addition to the information listed above.

**Inclusion Criteria:** The study population for this study includes people who: 1) are at least 18 years old, 2) are English speakers, 3) have perfect or corrected normal vision (contact lenses are supported). Participants should not have any symptom or known recent exposure to COVID-19.

**Who can I talk to?**
If you have questions, concerns, or complaints, or think the research has hurt you, please send an email to Feiyu Lu at feiyulu@vt.edu or talk to the research team directly at Room 160 (Sandbox), Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

**How many people will be studied?**
We plan to include about 12 people in this research study.

**What happens if I say yes, I want to be in this research?**
You will become a participant in this research. You will be invited to try out the AR system we developed, and give us your feedback about ways to interact with data and information in AR. The study will take place in the Sandbox Studio (Room 160, Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060).

**What are the experimental procedures?**
The experiment can be divided into five steps. In the first steps, you will be welcomed upon arrival. We will introduce briefly what our study is about, the motivations behind it, and explain to you any question you might have. Second, you will be asked to read and sign the consent form. Third, you will be asked to fill out a pre-study questionnaire to collect demographic information and prior experience with Augmented Reality. After completing these initial steps, fourth, you will run through a simple calibration process on the Magic Leap One Augmented Reality headset. Fourth, you will be asked to start using our application while sitting in front of a desktop computer or walking indoor. The application will display 4 to 6 pieces of virtual content (which contains random information such as artificial weather condition, to-do lists and calendar events) for you to glance at during the study. The application will place the virtual content in different scale and angles at the periphery. You will be asked to identify information in the virtual content displayed in AR, or information in the real-world environment behind the virtual content. During these use sessions, the application will record your head positions, orientations, gaze directions, and task completion time during use sessions. These
Consent to Take Part in a Research Study

sessions will also be video and audio recorded. Fifth, you will be asked to complete some post-study questionnaires, together with an interview asking about your thoughts on each of the conditions you experienced. The whole experiment will take roughly 20 minutes.

**What is the total duration of this study?**
The duration of an individual subject's participation will be 20 minutes.

**What happens if I say yes, but I change my mind later?**
You can leave the research at any time, for any reason, and it will not be held against you.

All the data collected to the point of withdrawal will be properly discarded and destroyed immediately.

**Is there any way being in this study could be bad for me? (Detailed Risks)**
Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment. Each participant will be accompanied by an experimenter alongside as a safety guarantee.

**What happens to the information collected for the research?**
We will make every effort to limit the use and disclosure of your personal information, including survey responses, diaries, and interview notes, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

**What else do I need to know?**
This research is being funded by *Office of Naval Research via University of California, Santa Barbara*

We will not offer to share your individual test results with you. You may accept or decline these results.

Your information and samples (both identifiable and de-identified) might be used to create products or to deliver services, including some that may be sold and/or make money for others. If this happens, there are no plans to tell you, or to pay you, or to give any compensation to you or your family.

There are no significant benefits that individual subjects might experience from participating in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset.

**Signature Block for Capable Adult**
Consent to Take Part in a Research Study

Your signature documents that you have read and understand the information outlined in this document and your permission to take part in this research. Please keep a digital signed copy of this form for your records.

Signature of subject ____________________________ Date ________________

Printed name of subject _______________________

Signature of person obtaining consent ________________ Date ________________

Printed name of person obtaining consent ____________________________

Virginia Tech Institutional Review Board Protocol No. 21-090
Approved March 11, 2021
B.1.2 Questionnaires
Demographic

Participants ID

Gender
- Male
- Female
- Other
- Prefer not to say

Age

Experience with AR
- 1 (Not)
Default Question Block

Which condition did you try just now?

- HUD-full app
- HUD-icon
- HeadGlance-full app
- HeadGlance-icon
- Empty

I found the interface very unobtrusive during walking.

Neither

Strongly disagree
Disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Agree
Strongly agree

I barely noticed the interface during walking.

Neither

Strongly disagree
Disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Agree
Strongly agree
I like walking around with this interface.

Could you state why you like or dislike the interface?

I found the interface annoying when I walk around with it.

If you found the interface annoying, could you state why?

I would like to use this interface in my everyday life to access information if the display was more like eye glasses.
Could you explain why you would like or not like to use the interface in everyday situations?

How often did the interface get in the way when you intended to look at the real-world or read the signs?

- Never
- Sometimes
- About half the time
- Most of the time
- Always

If the interface did get in the way, could you state in what situations it happened?

**Condition2**

Which condition did you try just now?

- HUD-full app
- HUD-icon
- HeadGlance-full app
I found the interface very unobtrusive during walking.

I barely noticed the interface during walking.

I like walking around with this interface.

Could you state why you like or dislike the interface?

Strongly disagree Disagree Somewhat disagree Neither Somewhat agree Agree Strongly agree

Strongly disagree Disagree Somewhat disagree Neither Somewhat agree Agree Strongly agree

Strongly disagree Disagree Somewhat disagree Neither Somewhat agree Agree Strongly agree

Strongly disagree Disagree Somewhat disagree Neither Somewhat agree Agree Strongly agree
I found the interface annoying when I walk around with it.

If you found the interface annoying, could you state why?

I would like to use this interface in my everyday life to access information if the display was more like eye glasses.

Could you explain why you would like or not like to use the interface in everyday situations?

How often did the interface get in the way when you intended to look at the real-world or read the signs?

About half the  Most of the
If the interface did get in the way, could you state in what situations it happened?

Condition 5

Which condition did you try just now?

- HUD-full app
- HUD-icon
- HeadGlance-full app
- HeadGlance-icon
- Empty

I found the interface very unobtrusive during walking.

Strongly disagree | Disagree | Somewhat agree nor disagree | Somewhat agree | Agree | Strongly agree
O | O | O | O | O | O

I barely noticed the interface during walking.
I like walking around with this interface.

Could you state why you like or dislike the interface?

I found the interface annoying when I walk around with it.

If you found the interface annoying, could you state why?
I would like to use this interface in my everyday life to access information if the display was more like eye glasses.

Could you explain why you would like or not like to use the interface in everyday situations?

How often did the interface get in the way when you intended to look at the real-world or read the signs?

If the interface did get in the way, could you state in what situations it happened?

**Condition 4**
Which condition did you try just now?

- HUD-full app
- HUD-icon
- HeadGlance-full app
- HeadGlance-icon
- Empty

I found the interface very unobtrusive during walking.

Neither

Strongly disagree

Disagree

Somewhat disagree

Agree

Strongly agree

I barely noticed the interface during walking.

Neither

Strongly disagree

Disagree

Somewhat disagree

Agree

Strongly agree

I like walking around with this interface.

Neither

Strongly disagree

Disagree

Somewhat disagree

Agree

Strongly agree
Could you state why you like or dislike the interface?

I found the interface annoying when I walk around with it.

If you found the interface annoying, could you state why?

I would like to use this interface in my everyday life to access information if the display was more like eye glasses.

Could you explain why you would like or not like to use the interface in everyday situations?
How often did the interface get in the way when you intended to look at the real-world or read the signs?

<table>
<thead>
<tr>
<th>Never</th>
<th>Sometimes</th>
<th>About half the time</th>
<th>Most of the time</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If the interface did get in the way, could you state in what situations it happened?

Condition3

Which condition did you try just now?

- HUD-full app
- HUD-icon
- HeadGlance-full app
- HeadGlance-icon
- Empty

I found the interface very unobtrusive during walking.
I barely noticed the interface during walking.

Neither
Strongly disagree Disagree Somewhat disagree Disagree Somewhat agree Agree Strongly agree

I like walking around with this interface.

Neither
Strongly disagree Disagree Somewhat disagree Disagree Somewhat agree Agree Strongly agree

Could you state why you like or dislike the interface?

I found the interface annoying when I walk around with it.

Neither
Strongly disagree Disagree Somewhat disagree Disagree Somewhat agree Agree Strongly agree
If you found the interface annoying, could you state why?

I would like to use this interface in my everyday life to access information if the display was more like eye glasses.

Could you explain why you would like or not like to use the interface in everyday situations?

How often did the interface get in the way when you intended to look at the real-world or read the signs?

If the interface did get in the way, could you state in what situations it happened?
B.2 Primary Study

B.2.1 Informed Consent Form
Title of research study: IRB #21-090: Identifying and resolving occlusion issues for Glanceable AR Interfaces

Principal Investigator: Dr. Doug A. Bowman  Contact: (540) 231-2058 or dbowman@vt.edu

Other study contact(s): Feiyu Lu  Contact: (540) 257-4562 or feiyulu@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

We invite you to take part in a research study about Augmented Reality (AR) user interfaces. The research will take around 90 minutes to complete.

What should I know about being in a research study?

- Someone will explain this research study to you
- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind
- Your decision will not be held against you
- You can ask all the questions you want before you decide.

What should I know about this research study? This research project is intended to compare a variety of methods for interacting with augmented reality (AR) systems. This research will help us understand the best ways to interact with data and information in AR and produce design guidelines for AR user interface design. You will be asked to wear a Magic Leap One AR headset and interact with several pieces of virtual content displayed in the headset while sitting or walking inside a room. Afterwards, you will be asked to fill out questionnaires about your experience (More detailed information can be found under “What are the experimental procedures?”). The research will take around 90 minutes to complete. There is risks that you might feel dizziness or nausea (More detailed information can be found under “Is there any way being in this study could be bad for me? (Detailed Risks)”). There are no significant benefits that individual subjects might experience from participating in this research. We cannot promise any benefits to others from your taking part in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset. Potential benefits to others also include the AR industry gaining a better understanding of occlusion issues in everyday AR uses. If you are a student at Virginia Tech, the
Consent to Take Part in a Research Study

decision whether to participate or not will have no effect on your academic performance or relationship with Virginia Tech.

**Detailed Information:** The following is more detailed information about this study in addition to the information listed above.

**Inclusion Criteria:** The study population for this study includes people who: 1) are at least 18 years old, 2) are English speakers, 3) have perfect or corrected normal vision (contact lenses are supported). Participants should not have any symptom or known recent exposure to COVID-19.

**Who can I talk to?**
If you have questions, concerns, or complaints, or think the research has hurt you, please send an email to Feiyu Lu at feiyulu@vt.edu or talk to the research team directly at Room 160 (Sandbox), Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060.

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

**How many people will be studied?**
We plan to include about 36 people in this research study.

**What happens if I say yes, I want to be in this research?**
You will become a participant in this research. You will be invited to try out the AR system we developed, and give us your feedback about ways to interact with data and information in AR. The study will take place in the Sandbox Studio (Room 160, Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060).

**What are the experimental procedures?**
The experiment can be divided into five steps. In the first steps, you will be welcomed upon arrival, and we will introduce briefly what our study is about, the motivations behind it, and explain to you any question you might have. Second, you will be asked to read and sign the consent form. Third, you will be asked to fill out a pre-study questionnaire to collect demographic information and prior experience with Augmented Reality. After completing these initial steps, fourth, you will run through a simple calibration process on the Magic Leap One Augmented Reality headset. Fourth, you will be asked to start using our application while sitting in front of a desktop computer or walking indoor. In both scenarios, the application will display 4 to 6 pieces of virtual content (which contains random information such as artificial weather condition, to-do lists and calendar events) for you to glance at during the study. The application will incorporate different interaction methods for you to toggle the visibility of the virtual content, including hands, eye, or head-based interactions. You will be asked to identify information in the virtual content displayed in AR, or information in the real-world environment behind the virtual content. During these use sessions, the application will record your
Consent to Take Part in a Research Study

head positions, orientations, gaze directions, and task completion time during use sessions. These sessions will also be video and audio recorded. Fifth, you will be asked to complete some post-study questionnaires, together with an interview asking about your thoughts on each of the conditions you experienced. The whole experiment will take roughly 90 minutes.

What is the total duration of this study?
The duration of an individual subject's participation will be 90 minutes.

What happens if I say yes, but I change my mind later?
You can leave the research at any time, for any reason, and it will not be held against you. All the data collected to the point of withdrawal will be properly discarded and destroyed immediately.

Is there any way being in this study could be bad for me? (Detailed Risks)
Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment. Each participant will be accompanied by an experimenter alongside as a safety guarantee.

What happens to the information collected for the research?
We will make every effort to limit the use and disclosure of your personal information, including survey responses, diaries, and interview notes, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

What else do I need to know?
This research is being funded by Office of Naval Research via University of California, Santa Barbara

We will not offer to share your individual test results with you. You may accept or decline these results.

Your information and samples (both identifiable and de-identified) might be used to create products or to deliver services, including some that may be sold and/or make money for others. If this happens, there are no plans to tell you, or to pay you, or to give any compensation to you or your family.

There are no significant benefits that individual subjects might experience from participating in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset.
Consent to Take Part in a Research Study

Signature Block for Capable Adult

Your signature documents that you have read and understand the information outlined in this
document and your permission to take part in this research. Please keep a digital signed copy of this
form for your records.

<table>
<thead>
<tr>
<th>Signature of subject</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Printed name of subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature of person obtaining consent</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Printed name of person obtaining consent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
B.2.2 Questionnaires
Default Question Block

Gender

- Male
- Female
- Non-binary / third gender
- Prefer not to say

Age


Are you:

- Right-handed
- Left-handed
- Ambidextrous
- Other
Occupation (if student, indicate graduate or undergraduate):


Major / Area of specialization (if student):


Rate your fatigue level:

1 (Very tired) 2 3 4 5 (Not tired at all)

How many times have you tried Augmented/Mixed reality ?

Never
Once or twice
3-10 times
More than 10 times
B.2.3 Semi-structured Interview Scriptss
## Technique 1

How do you like this technique for accessing information in the widget?

<table>
<thead>
<tr>
<th>Extremely bad</th>
<th>Somewhat bad</th>
<th>Neither good nor bad</th>
<th>Somewhat good</th>
<th>Extremely good</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How do you like this technique for accessing information in the sign?

<table>
<thead>
<tr>
<th>Extremely bad</th>
<th>Somewhat bad</th>
<th>Neither good nor bad</th>
<th>Somewhat good</th>
<th>Extremely good</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I think that I would like to use the technique frequently.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I found the technique unnecessarily complex.

<table>
<thead>
<tr>
<th>Strongly</th>
<th>Somewhat</th>
<th>Neither agree</th>
<th>Somewhat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I thought the technique was easy to use.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I think that I would need the support of a technical person to be able to use this technique.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I would imagine that most people would learn to use this technique very quickly.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I found the technique very cumbersome to use.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree
I felt very confident using the technique.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

I needed to learn a lot of things before I could get going with this technique.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

How mentally demanding was the task?

- Very Low: ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ | Very High

How physically demanding was the task?

- Very Low: ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ | Very High

How hurried or rushed was the pace of the task?

- Very Low: ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ | Very High
How successful were you in accomplishing what you were asked to do?

Perfect   ○ ○ ○ ○ ○ ○ ○ ○   Failure

How hard did you have to work to accomplish your level of performance?

Very Low   ○ ○ ○ ○ ○ ○ ○ ○   Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low   ○ ○ ○ ○ ○ ○ ○ ○   Very High

On a scale from 1-7, what was your overall emotion/impression during the task?

1 (I hated it. It felt terribly awkward)  2  3  4  5  6  7 (I enjoyed it. It felt comfortable)

In front of whom would you feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.
In which locations do you think you would feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

☐ When alone
☐ In front of my partner
☐ In front of friends
☐ In front of family
☐ In front of colleagues
☐ In front of strangers

In which locations do you think you would feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

☐ At home
☐ On the sidewalk
☐ In a pub, café, or a restaurant
☐ In a shop
☐ In a museum
☐ As a passenger on a bus or train
☐ At my workplace

Could you comment on what you like and dislike about this technique?
Technique 2

How do you like this technique for accessing information in the widget?

- Extremely bad
- Somewhat bad
- Neither good nor bad
- Somewhat good
- Extremely good

How do you like this technique for accessing information in the sign?

- Extremely bad
- Somewhat bad
- Neither good nor bad
- Somewhat good
- Extremely good

I think that I would like to use the technique frequently.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I found the technique unnecessarily complex.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I thought the technique was easy to use.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
I think that I would need the support of a technical person to be able to use this technique.

I would imagine that most people would learn to use this technique very quickly.

I found the technique very cumbersome to use.

I felt very confident using the technique.
I needed to learn a lot of things before I could get going with this technique.

How mentally demanding was the task?

Very Low  ○ ○ ○ ○ ○ ○ ○ ○ Very High

How physically demanding was the task?

Very Low  ○ ○ ○ ○ ○ ○ ○ ○ Very High

How hurried or rushed was the pace of the task?

Very Low  ○ ○ ○ ○ ○ ○ ○ ○ Very High

How successful were you in accomplishing what you were asked to do?

Perfect  ○ ○ ○ ○ ○ ○ ○ ○ Failure

How hard did you have to work to accomplish your level of
performance?

Very Low  ◯  ◯  ◯  ◯  ◯  ◯  ◯  ◯  Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low  ◯  ◯  ◯  ◯  ◯  ◯  ◯  ◯  Very High

On a scale from 1-7, what was your overall emotion/impression during the task?

1 (I hated it. It felt terribly awkward)  ◯  ◯  ◯  ◯  ◯  ◯  ◯  ◯  7 (I enjoyed it. It felt comfortable)

In front of whom would you feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

☐ When alone
☐ In front of my partner
☐ In front of friends
☐ In front of family
☐ In front of colleagues
☐ In front of strangers
In which locations do you think you would feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

- [ ] At home
- [ ] On the sidewalk
- [ ] In a pub, café, or a restaurant
- [ ] In a shop
- [ ] In a museum
- [ ] As a passenger on a bus or train
- [ ] At my workplace

Could you comment on what you like and dislike about this technique?

Technique 3

How do you like this technique for accessing information in the widget?

- [ ] Extremely bad
- [ ] Somewhat bad
- [ ] Neither good nor bad
- [ ] Somewhat good
- [ ] Extremely good
How do you like this technique for accessing information in the sign?

- Extremely bad
- Somewhat bad
- Neither good nor bad
- Somewhat good
- Extremely good

I think that I would like to use the technique frequently.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I found the technique unnecessarily complex.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I thought the technique was easy to use.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I think that I would need the support of a technical person to be able to use this technique.

- Strongly disagree
- Somewhat disagree
- Neither agree
- Somewhat agree
- Strongly agree
I would imagine that most people would learn to use this technique very quickly.

I found the technique very cumbersome to use.

I felt very confident using the technique.

I needed to learn a lot of things before I could get going with this technique.
How mentally demanding was the task?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ Very High

How physically demanding was the task?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ Very High

How hurried or rushed was the pace of the task?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ Very High

How successful were you in accomplishing what you were asked to do?

Perfect ○ ○ ○ ○ ○ ○ ○ ○ Failure

How hard did you have to work to accomplish your level of performance?

Very Low ○ ○ ○ ○ ○ ○ ○ ○ Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?
On a scale from 1-7, what was your overall emotion/impression during the task?

1 (I hated it. It felt terribly awkward)

2

3

4

5

6

7 (I enjoyed it. It felt comfortable)

In front of whom would you feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

- When alone
- In front of my partner
- In front of friends
- In front of family
- In front of colleagues
- In front of strangers

In which locations do you think you would feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

- At home
Could you comment on what you like and dislike about this technique?

Technique 4

How do you like this technique for accessing information in the widget?

- [ ] Extremely bad
- [ ] Somewhat bad
- [ ] Neither good nor bad
- [ ] Somewhat good
- [ ] Extremely good

How do you like this technique for accessing information in the sign?

- [ ] Extremely bad
- [ ] Somewhat bad
- [ ] Neither good nor bad
- [ ] Somewhat good
- [ ] Extremely good
I think that I would like to use the technique frequently.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I found the technique unnecessarily complex.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I thought the technique was easy to use.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I think that I would need the support of a technical person to be able to use this technique.

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

I would imagine that most people would learn to use this technique very quickly.

- Strongly disagree
- Somewhat disagree
- Neither agree
- Somewhat agree
- Strongly agree
I found the technique very cumbersome to use.

I felt very confident using the technique.

I needed to learn a lot of things before I could get going with this technique.

How mentally demanding was the task?

How physically demanding was the task?
How hurried or rushed was the pace of the task?

Very Low 〇 〇 〇 〇 〇 〇 〇 〇  Very High

How successful were you in accomplishing what you were asked to do?

Perfect 〇 〇 〇 〇 〇 〇 〇 〇  Failure

How hard did you have to work to accomplish your level of performance?

Very Low 〇 〇 〇 〇 〇 〇 〇 〇  Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low 〇 〇 〇 〇 〇 〇 〇 〇  Very High

On a scale from 1-7, what was your overall emotion/impression during the task?

1 (I hated it. It felt 7 (I enjoyed
In front of whom would you feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

- When alone
- In front of my partner
- In front of friends
- In front of family
- In front of colleagues
- In front of strangers

In which locations do you think you would feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

- At home
- On the sidewalk
- In a pub, café, or a restaurant
- In a shop
- In a museum
- As a passenger on a bus or train
- At my workplace
Could you comment on what you like and dislike about this technique?

Technique 5

How do you like this technique for accessing information in the widget?

Extremely bad  Somewhat bad  Neither good nor bad  Somewhat good  Extremely good

How do you like this technique for accessing information in the sign?

Extremely bad  Somewhat bad  Neither good nor bad  Somewhat good  Extremely good

I think that I would like to use the technique frequently.

Strongly disagree  Somewhat disagree  Neither agree nor disagree  Somewhat agree  Strongly agree
<table>
<thead>
<tr>
<th>I found the technique unnecessarily complex.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
</tr>
<tr>
<td><img src="circle" alt="Circle" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I thought the technique was easy to use.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
</tr>
<tr>
<td><img src="circle" alt="Circle" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I think that I would need the support of a technical person to be able to use this technique.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
</tr>
<tr>
<td><img src="circle" alt="Circle" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I would imagine that most people would learn to use this technique very quickly.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
</tr>
<tr>
<td><img src="circle" alt="Circle" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I found the technique very cumbersome to use.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
</tr>
<tr>
<td><img src="circle" alt="Circle" /></td>
</tr>
</tbody>
</table>
I felt very confident using the technique.

I needed to learn a lot of things before I could get going with this technique.

How mentally demanding was the task?

Very Low

How physically demanding was the task?

Very Low

How hurried or rushed was the pace of the task?

Very Low
How successful were you in accomplishing what you were asked to do?

Perfect ○ ○ ○ ○ ○ ○ ○ ○ Failure

How hard did you have to work to accomplish your level of performance?

Very Low ○ ○ ○ ○ ○ ○ ○ Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low ○ ○ ○ ○ ○ ○ ○ Very High

On a scale from 1-7, what was your overall emotion/impression during the task?

1 (I hated it. It felt terribly awkward) 2 3 4 5 6 7 (I enjoyed it. It felt comfortable)

In front of whom would you feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply
from the list below.

☐ When alone
☐ In front of my partner
☐ In front of friends
☐ In front of family
☐ In front of colleagues
☐ In front of strangers

In which locations do you think you would feel comfortable using this technique (ignoring the form factor of the display)? Select all the items that apply from the list below.

☐ At home
☐ On the sidewalk
☐ In a pub, café, or a restaurant
☐ In a shop
☐ In a museum
☐ As a passenger on a bus or train
☐ At my workplace

Could you comment on what you like and dislike about this technique?
Rank the techniques

Please rank these interface based on your own preferences for the walking scenario.

<table>
<thead>
<tr>
<th>Items</th>
<th>Your ranking (best to worst)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand Overlap</td>
<td></td>
</tr>
<tr>
<td>Head Depth</td>
<td></td>
</tr>
<tr>
<td>Dwell</td>
<td></td>
</tr>
<tr>
<td>Blink</td>
<td></td>
</tr>
<tr>
<td>Fixation Glance</td>
<td></td>
</tr>
</tbody>
</table>

Experimenter Logs

Condition

- Sitting
- Walking

Order of testing

https://virginiatech.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurvey?surveyId=SV_37VRi4tiyWrink5Qa&contextLibraryId=UR_dIZWdhZSLge4B
Particpants ID
B.3 IRB Approval Letter
MEMORANDUM

DATE: March 11, 2021
TO: Douglas Andrew Bowman, Feiyu Lu, Shakiba Davari-Najafabadi
FROM: Virginia Tech Institutional Review Board (FWA00000572, expires October 29, 2024)

PROTOCOL TITLE: Identifying and Resolving Occlusion Issues for Glanceable AR Interfaces
IRB NUMBER: 21-090

Effective March 11, 2021, the Virginia Tech Institution Review Board (IRB) approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at: https://secure.research.vt.edu/external/irb/responsibilities.htm
(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Approved As: Expedited, under 45 CFR 46.110 category(ies) 4,6,7
Protocol Approval Date: March 11, 2021
Progress Review Date: March 11, 2022

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.
<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.
B.4 Video Demonstrations

Video demonstrations of this research is available both online here and in the supplementary material files that are attached to this dissertation. The presentation recording of this paper is available here.
Appendix C

Supporting Spatial Transitions with AR Interfaces

C.1 Modelling User Interactions During Spatial Transition

C.1.1 Informed Consent Form
Title of research study: **IRB #21-843: Exploring Transition Strategies of Glanceable Information in head-Worn Augmented Reality**

Principal Investigator: Dr. Doug A. Bowman  Contact: (540) 231-2058 or dbowman@vt.edu

Other study contact(s): Feiyu Lu  Contact: (540) 257-4562 or feiyulu@vt.edu; Leonardo Pavanatto: Contact: lpavanat@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

We invite you to take part in a research study about Augmented Reality (AR) user interfaces. The research will take around 45 minutes to complete.

**What should I know about being in a research study?**
- Someone will explain this research study to you
- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind
- Your decision will not be held against you
- You can ask all the questions you want before you decide.

**What should I know about this research study?** This research project is intended to compare a variety of methods for interacting with augmented reality (AR) systems. This research will help us understand the best ways to interact with data and information in AR and produce design guidelines for AR user interface design. You will be asked to wear a Magic Leap One AR headset and interact with several pieces of virtual content displayed in the headset while sitting or walking inside a room. Afterwards, you will be asked to fill out questionnaires about your experience (More detailed information can be found under “What are the experimental procedures?”). The research will take around 45 minutes to complete. There is risks that you might feel dizziness or nausea (More detailed information can be found under “Is there any way being in this study could be bad for me? (Detailed Risks)”). There are no significant benefits that individual subjects might experience from participating in this research. We cannot promise any benefits to others from your taking part in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset. Potential benefits to others also include the AR industry gaining a better understanding of usability issues of AR interfaces to support everyday uses. If you are a student at
Consent to Take Part in a Research Study

Virginia Tech, the decision whether to participate or not will have no effect on your academic performance or relationship with Virginia Tech.

**Detailed Information:** The following is more detailed information about this study in addition to the information listed above.

**Inclusion Criteria:** The study population for this study includes people who: 1) are at least 18 years old, 2) are English speakers, 3) have perfect or corrected normal vision (contact lenses are supported).

**Who can I talk to?**
If you have questions, concerns, or complaints, or think the research has hurt you, please send an email to Feiyu Lu at feiyulu@vt.edu or talk to the research team directly at Room 160 (Sandbox), Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

**How many people will be studied?**
We plan to include about 45 people in this research study.

**What happens if I say yes, I want to be in this research?**
You will become a participant in this research. You will be invited to try out the AR system we developed, and give us your feedback about ways to interact with data and information in AR. The study will take place in the Sandbox Studio (Room 160, Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060).

**What are the experimental procedures?**
The experiment can be divided into five steps. In the first steps, you will be welcomed upon arrival. We will introduce briefly what our study is about, the motivations behind it, and explain to you any question you might have. Second, you will be asked to read and sign the consent form. Third, you will be asked to fill out a pre-study questionnaire to collect demographic information and prior experience with Augmented Reality. After completing these initial steps, fourth, you will run through a simple calibration process on the Magic Leap One Augmented Reality headset. Fourth, you will be asked to start using our application while moving indoor. The application will display up to 28 pieces of virtual content which contains random icon and computer-generated phrases. Some of these content will be located in the physical world while some of them will be inside a menu for you to access. You will be asked to search for some virtual content displayed in AR and identify the content in it. During these use sessions, the application will record your head positions, orientations, gaze directions, and task completion time during use sessions. These sessions will also be audio recorded. Fifth, you will be asked to complete some post-study questionnaires, together with an interview asking about your
thoughts on the experience. The interview will be audio-recorded. The whole experiment will take roughly 45 minutes.

**What is the total duration of this study?**
The duration of an individual subject's participation will be 45 minutes.

**What happens if I say yes, but I change my mind later?**
You can leave the research at any time, for any reason, and it will not be held against you.
All the data collected to the point of withdrawal will be properly discarded and destroyed immediately.

**Is there any way being in this study could be bad for me? (Detailed Risks)**
Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment. Each participant will be accompanied by an experimenter alongside as a safety guarantee.

**What happens to the information collected for the research?**
We will make every effort to limit the use and disclosure of your personal information, including survey responses, diaries, and interview notes, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

**What else do I need to know?**
This research is being funded by *Office of Naval Research via University of California, Santa Barbara*

We will not offer to share your individual test results with you. You may accept or decline these results.

Your information and samples (both identifiable and de-identified) might be used to create products or to deliver services, including some that may be sold and/or make money for others. If this happens, there are no plans to tell you, or to pay you, or to give any compensation to you or your family.

There are no significant benefits that individual subjects might experience from participating in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset.

**Signature Block for Capable Adult**
## Consent to Take Part in a Research Study

Your signature documents that you have read and understand the information outlined in this document and your permission to take part in this research. Please keep a digital signed copy of this form for your records.

<table>
<thead>
<tr>
<th>Signature of subject</th>
<th>Date</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Printed name of subject</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Signature of person obtaining consent</th>
<th>Date</th>
</tr>
</thead>
</table>

| Printed name of person obtaining consent |   |

---

Virginia Tech Institutional Review Board Protocol No. 21-843
Approved October 29, 2021
C.1.2 Questionnaires & Interviews
Demographic

Gender

- Male
- Female
- Non-binary / third gender
- Prefer not to say

Age

Are you

- Right-handed
- Left-handed
- Ambidextrous
- Other
Occupation (if student, indicate graduate or undergraduate):

Major / Area of specialization (if student):

Please rate your fatigue level:

1 (Not tired at all) 2 3 4 5 (Very tired)

Please rate your experience with Augmented Reality:

1 (Not experienced at all) 2 3 4 5 (Very experienced)

Condition 1

Which condition did you just try?

None
Considering your experience of completing one single trial of this condition, please answer the following questions.

I found it easy to locate the widgets I need.

- Strongly Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

How mentally demanding was the task?

- Very Low
- Very High

How physically demanding was the task?

- Very Low
- Very High

How successful were you in accomplishing what you were asked to do?
How hard did you have to work to accomplish your level of performance?

Very Low  ○ ○ ○ ○ ○ ○ ○ ○  Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low  ○ ○ ○ ○ ○ ○ ○ ○  Very High

Could you state what you like and dislike about this condition?

Condition 2

Which condition did you just try?

○ None
○ Some
○ All
Considering your experience of completing one single trial of this condition, please answer the following questions.

I found it easy to locate the widgets I need.

- [ ] Strongly Agree
- [ ] Somewhat agree
- [ ] Neither agree nor disagree
- [ ] Somewhat disagree
- [ ] Strongly disagree

How mentally demanding was the task?

Very Low: [ ] [ ] [ ] [ ] [ ] [ ] [ ] Very High

How physically demanding was the task?

Very Low: [ ] [ ] [ ] [ ] [ ] [ ] [ ] Very High

How successful were you in accomplishing what you were asked to do?

Perfect: [ ] [ ] [ ] [ ] [ ] [ ] [ ] Failure
How hard did you have to work to accomplish your level of performance?

Very Low 〇 〇 〇 〇 〇 〇 〇 〇 Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low 〇 〇 〇 〇 〇 〇 〇 〇 Very High

Could you state what you like and dislike about this condition?

Condition 3

Which condition did you just try?

〇 None
〇 Some
〇 All

Considering your experience of completing one single trial of this
condition, please answer the following questions

I found it easy to locate the widgets I need.

- Strongly Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

How mentally demanding was the task?

Very Low

Very High

How physically demanding was the task?

Very Low

Very High

How successful were you in accomplishing what you were asked to do?

Perfect

Failure

How hard did you have to work to accomplish your level of
performance?

Very Low   ○ ○ ○ ○ ○ ○ ○   Very High

How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low   ○ ○ ○ ○ ○ ○ ○   Very High

Could you state what you like and dislike about this condition?

End of Survey

Considering your experience of completing one single trial of each condition, please rank the three conditions based on your own preference:

<table>
<thead>
<tr>
<th>Items</th>
<th>Click to write Group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Some</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td></td>
</tr>
</tbody>
</table>
Please enter the participant ID:

Please enter the order of testing:

Powered by Qualtrics
C.2  Evaluating Realistic Transitions In-Situ

C.2.1  Informed Consent Form
Title of research study: IRB #21-843: Exploring Transition Strategies of Glanceable Information in head-Worn Augmented Reality

Principal Investigator: Dr. Doug A. Bowman  Contact: (540) 231-2058 or dbowman@vt.edu

Other study contact(s): Feiyu Lu  Contact: (540) 257-4562 or feiyulu@vt.edu; Leonardo Pavanatto  Contact: lpavanat@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

We invite you to take part in a research study about Augmented Reality (AR) user interfaces. The research will take around 80 minutes to complete. Participant will be compensated by either Virginia Tech Ut-Prosim service credits or $15 via Amazon e-gift card, the option is up to the participant.

What should I know about being in a research study?

- Someone will explain this research study to you
- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind
- Your decision will not be held against you
- You can ask all the questions you want before you decide.

What should I know about this research study? This research project is intended to compare a variety of methods for interacting with augmented reality (AR) systems. This research will help us understand the best ways to interact with data and information in AR and produce design guidelines for AR user interface design. You will be asked to wear a Magic Leap One AR headset and interact with several pieces of virtual content displayed in the headset while sitting or walking inside a room. Afterwards, you will be asked to fill out questionnaires about your experience (More detailed information can be found under “What are the experimental procedures?”). The research will take around 90 minutes to complete. There is risks that you might feel dizziness or nausea (More detailed information can be found under “Is there any way being in this study could be bad for me? (Detailed Risks)”). There are no significant benefits that individual subjects might experience from participating in this research. We cannot promise any benefits to others from your taking part in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset. Potential benefits to others also include the AR industry gaining a better understanding of usability issues of AR interfaces to support everyday uses. If you are a student at
Consent to Take Part in a Research Study

Virginia Tech, the decision whether to participate or not will have no effect on your academic performance or relationship with Virginia Tech.

**Detailed Information:** The following is more detailed information about this study in addition to the information listed above.

**Inclusion Criteria:** The study population for this study includes people who: 1) are at least 18 years old, 2) are English speakers, 3) have perfect or corrected normal vision (contact lenses are supported).

**Who can I talk to?**
If you have questions, concerns, or complaints, or think the research has hurt you, please send an email to Feiyu Lu at feiyulu@vt.edu or talk to the research team directly at Room 160 (Sandbox), Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

**How many people will be studied?**
We plan to include about 30 people in this research study.

**What happens if I say yes, I want to be in this research?**
You will become a participant in this research. You will be invited to try out the AR system we developed, and give us your feedback about ways to interact with data and information in AR. The study will take place in 2202 Kraft Drive, Blacksburg, VA 24060.

**What are the experimental procedures?**
The experiment can be divided into five steps. In the first steps, you will be welcomed upon arrival, and we will introduce briefly what our study is about, the motivations behind it, and explain to you any question you might have. Second, you will be asked to read and sign the consent form. Third, you will be asked to fill out a pre-study questionnaire to collect demographic information and prior experience with Augmented Reality. After completing these initial steps, fourth, you will run through a simple calibration process on the Magic Leap One Augmented Reality headset. Fourth, you will be asked to start using our application while moving between 4-5 locations in an indoor environment. The application will display 15-30 pieces of virtual content which contains random information such as icons and computer-generated phrases. Some of these content will be located in the physical world while some of them will be inside a menu for you to access. You will be asked to search for some virtual content displayed in AR and identify the content in it while moving between these locations with different levels of system automation and user input. You will be asked to identify information in the virtual content displayed in AR, or information in the real-world environment behind the virtual content. During these use sessions, the application will record your head positions, orientations, gaze...
Consent to Take Part in a Research Study

directions, and task completion time during use sessions. These sessions will be audio-recorded. Fifth, you will be asked to complete some post-study questionnaires, together with an interview asking about your thoughts on each of the conditions you experienced. The interview will be audio-recorded. The whole experiment will take roughly 80 minutes. Participant will be compensated by either Virginia Tech Ut-Prosim service credits or $15 via Amazon e-gift card, the option is up to the participant.

What is the total duration of this study?
The duration of an individual subject's participation will be 80 minutes.

What happens if I say yes, but I change my mind later?
You can leave the research at any time, for any reason, and it will not be held against you. All the data collected to the point of withdrawal will be properly discarded and destroyed immediately.

Is there any way being in this study could be bad for me? (Detailed Risks)
Using AR technology can produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation. During tasks involving physical movement, there is also some risk that you could collide with obstacles in the physical environment. Each participant will be accompanied by an experimenter alongside as a safety guarantee.

What happens to the information collected for the research?
We will make every effort to limit the use and disclosure of your personal information, including survey responses, diaries, and interview notes, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

What else do I need to know?
This research is being funded by Office of Naval Research via University of California, Santa Barbara. We will not offer to share your individual test results with you. You may accept or decline these results.

Your information and samples (both identifiable and de-identified) might be used to create products or to deliver services, including some that may be sold and/or make money for others. If this happens, there are no plans to tell you, or to pay you, or to give any compensation to you or your family.

There are no significant benefits that individual subjects might experience from participating in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset.
Consent to Take Part in a Research Study

Signature Block for Capable Adult

Your signature documents that you have read and understand the information outlined in this document and your permission to take part in this research. Please keep a digital signed copy of this form for your records.

<table>
<thead>
<tr>
<th>Signature of subject</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Printed name of subject</td>
<td></td>
</tr>
<tr>
<td>Signature of person obtaining consent</td>
<td>Date</td>
</tr>
<tr>
<td>Printed name of person obtaining consent</td>
<td></td>
</tr>
</tbody>
</table>
C.2.2 Questionnaires
Transition Study Questionnaire

* Indicates required question

1. Gender

* Mark only one oval.

- [ ] Male
- [ ] Female
- [ ] Non-binary
- [ ] Prefer not to say

2. Age

---

3. Are you:

* Mark only one oval.

- [ ] Right-handed
- [ ] Left-handed
- [ ] Ambidextrous
- [ ] Other: ____________________
4. Please rate your fatigue level

*Mark only one oval.*

1 2 3 4 5

Not ☐ ☐ ☐ ☐ ☐ Very tired

5. Please rate your experience with Augmented Reality (AR):

*Mark only one oval.*

1 2 3 4 5

No I ☐ ☐ ☐ ☐ ☐ Very Experienced

After trying interface 3

6. Overall, how difficult or easy did you find the task? *

*Mark only one oval.*

1 2 3 4 5

Very ☐ ☐ ☐ ☐ ☐ Very Easy
7. Which condition did you just try? *

Mark only one oval.

- None
- Some
- All

8. I think that I would like to use this interface frequently. *

Mark only one oval.

1  2  3  4  5

Stro ○ ○ ○ ○ ○ Strongly Agree

9. I found this interface unnecessary complex. *

Mark only one oval.

1  2  3  4  5

Stro ○ ○ ○ ○ ○ Strongly Agree

10. I thought the interface was easy to use. *

Mark only one oval.

1  2  3  4  5

Stro ○ ○ ○ ○ ○ Strongly Agree
11. I think that I would need the support from a technical person to be able to use this interface. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

12. I would imagine that most people would learn to use the interface very quickly. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

13. I found the interface very cumbersome to use. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

14. I felt very confident using the interface. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree
15. I needed to learn a lot of things before I could get going with this interface. *

Mark only one oval.

1 2 3 4 5

Stro ☐ ☐ ☐ ☐ ☐ Strongly Agree

16. How mentally demanding was the task? *

Mark only one oval.

1 2 3 4 5

Very ☐ ☐ ☐ ☐ ☐ Very High

17. How physically demanding was the task *

Mark only one oval.

1 2 3 4 5

Very ☐ ☐ ☐ ☐ ☐ Very High

18. I felt that I was in full control over the virtual applications. *

Mark only one oval.

1 2 3 4 5

Stro ☐ ☐ ☐ ☐ ☐ Strongly Agree
19. I felt that the virtual applications moved without my intent. *

Mark only one oval.

1 2 3 4 5

Strongly Agree

20. I felt that the automated applications were accurate when I arrive at a new location. Please ignore the question for the None condition

Mark only one oval.

1 2 3 4 5

Strongly Agree

21. I felt that I was responsible for the success of the tasks. *

Mark only one oval.

1 2 3 4 5

Strongly Agree

22. I found the controls for the virtual applications satisfactory for the completion of the task. *

Mark only one oval.

1 2 3 4 5

Strongly Agree
23. Please comment on what you like/dislike about this interface.


After trying interface 2

24. Overall, how difficult or easy did you find the task? *

Mark only one oval.

1 2 3 4 5

Very □ □ □ □ □ Very Easy

25. Which condition did you just try? *

Mark only one oval.

□ None
□ Some
□ All
26. I think that I would like to use this interface frequently. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○○ Strongly Agree

27. I found this interface unnecessary complex. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○○ Strongly Agree

28. I thought the interface was easy to use. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○○ Strongly Agree

29. I think that I would need the support from a technical person to be able to use this interface. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○○ Strongly Agree
30. I would imagine that most people would learn to use the interface very quickly. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

31. I found the interface very cumbersome to use. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

32. I felt very confident using the interface. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

33. I needed to learn a lot of things before I could get going with this interface. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree
34. How mentally demanding was the task? *

*Mark only one oval.*

1 2 3 4 5

Very ☐ ☐ ☐ ☐ ☐ Very High

35. How physically demanding was the task *

*Mark only one oval.*

1 2 3 4 5

Very ☐ ☐ ☐ ☐ ☐ Very High

36. I felt that I was in full control over the virtual applications. *

*Mark only one oval.*

1 2 3 4 5

Stro ☐ ☐ ☐ ☐ ☐ Strongly Agree

37. I felt that the virtual applications moved without my intent. *

*Mark only one oval.*

1 2 3 4 5

Stro ☐ ☐ ☐ ☐ ☐ Strongly Agree
38. I felt that the automated applications were accurate when I arrive at a new location. Please ignore the question for the None condition

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

39. I felt that I was responsible for the success of the tasks. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

40. I found the controls for the virtual applications satisfactory for the completion of the task. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

41. Please comment on what you like/dislike about this interface.

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________

____________________________________________________
After trying interface 1

42. Overall, how difficult or easy did you find the task? *

*Mark only one oval.*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very</td>
<td></td>
<td></td>
<td></td>
<td>Very Easy</td>
</tr>
</tbody>
</table>

43. Which condition did you just try? *

*Mark only one oval.*

- None
- Some
- All

44. I think that I would like to use this interface frequently. *

*Mark only one oval.*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly</td>
<td></td>
<td></td>
<td></td>
<td>Agree</td>
</tr>
</tbody>
</table>

Strongly Disagree
45. I found this interface unnecessary complex. *

Mark only one oval.

1  2  3  4  5

Strongly Agree

46. I thought the interface was easy to use. *

Mark only one oval.

1  2  3  4  5

Strongly Agree

47. I think that I would need the support from a technical person to be able to use this interface. *

Mark only one oval.

1  2  3  4  5

Strongly Agree

48. I would imagine that most people would learn to use the interface very quickly. *

Mark only one oval.

1  2  3  4  5

Strongly Agree
49. I found the interface very cumbersome to use. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

50. I felt very confident using the interface. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

51. I needed to learn a lot of things before I could get going with this interface. *

*Mark only one oval.*

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

52. How mentally demanding was the task? *

*Mark only one oval.*

1 2 3 4 5

Very ○ ○ ○ ○ ○ Very High
53. How physically demanding was the task *

Mark only one oval.

1 2 3 4 5

Very ○ ○ ○ ○ ○ Very High

54. I felt that I was in full control over the virtual applications. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

55. I felt that the virtual applications moved without my intent. *

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree

56. I felt that the automated applications were accurate when I arrive at a new location.
Please ignore the question for the None condition

Mark only one oval.

1 2 3 4 5

Stro ○ ○ ○ ○ ○ Strongly Agree
57. I felt that I was responsible for the success of the tasks. *

Mark only one oval.

1  2  3  4  5

Stro ○ ○ ○ ○ ○ Strongly Agree

58. I found the controls for the virtual applications satisfactory for the completion of the task. *

Mark only one oval.

1  2  3  4  5

Stro ○ ○ ○ ○ ○ Strongly Agree

59. Please comment on what you like/dislike about this interface.

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________

Section 6
60. Please rank the four interfaces you just tried: *

*Mark only one oval per row.*

<table>
<thead>
<tr>
<th></th>
<th>Most Favorite</th>
<th>2nd Favorite</th>
<th>3rd Favorite</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>None</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Some</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

61. PID

62. Order of testing:

---

This content is neither created nor endorsed by Google.

Google Forms
C.2.3 Semi-structured Interview Scriptss
Exploring Transition Strategies of Glanceable Information in Head-Worn Augmented Reality

Semi-structured Interview topics/questions:

1. What did you like and dislike about the three conditions?
2. In a contextual situation like XX when XX, how did you feel like using the three interfaces? What are the pros and cons? Which interfaces did you prefer more?
3. In what ways do you feel the system could be further improved?
4. Could you elaborate more on the phrase “XX” you just mentioned?
5. Could you elaborate more on why you think the interface was “XX”?
C.3 IRB Approval Letter
Effective October 18, 2022, the Virginia Tech Institution Review Board (IRB) approved the Amendment request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at: https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

**PROTOCOL INFORMATION:**

- Approved As: Expedited, under 45 CFR 46.110 category(ies) 4,6,7
- Protocol Approval Date: October 17, 2022
- Progress Review Date: October 16, 2023

**ASSOCIATED FUNDING:**

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.
SPECIAL INSTRUCTIONS:
This amendment, submitted October 18, 2022, updates research protocol, recruitment materials, and consent form.

<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/19/2021</td>
<td>PET47EVN</td>
<td>University of California, Santa Barbara (Title: View management and user interface optimization for wide-area mobile augmented real)</td>
<td>Compared on 10/19/2021</td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.
C.4 Video Demonstrations

Video demonstrations of this research is available both online here and in the supplementary material files that are attached to this dissertation.
Appendix D

Real-world Deployments of Glanceable AR

D.1 Surveying Public Opinions About a Glanceable AR Video Prototype

D.1.1 Informed Consent Form & Questionnaire
Consent

Title of research study: IRB #20-519: Glanceable AR: Evaluating User Perceptions on Future Augmented Reality Interfaces

Principal Investigator: Dr. Doug A. Bowman Contact: (540) 231-2058 or dbowman@vt.edu

Other study contact(s): Feiyu Lu Contact: (540) 257-4562 or feiyulu@vt.edu

You are invited to participate in an online survey on an augmented reality video prototype we developed. This is a research project being conducted by Feiyu Lu, a student at Virginia Tech. The study should take approximately 30 minutes to complete.

What is the purpose of this study?

This research project is intended to gain feedback on a video prototype we built, which showcases future use cases of augmented reality displays in everyday scenarios. This research will help us understand user perceptions on usages of augmented reality to assist everyday activities and produce design guidelines for futuristic augmented reality interfaces that can be well integrated into the background of our lives.

How many people will be studied?

We plan to include up to 500 people in this research study.
What should I know about participating in a research study?

- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind with no penalty
- Your decision will not be held against you
- You can ask all the questions you want before you decide

What happens if I say yes, I want to be in this research?

You will become a participant in this research. You will be invited to watch a video and complete an online survey, which will take no longer than 30 minutes. If you are interested in participating in an additional interview via Zoom, you are welcomed to leave your email address at the end of the survey. We will contact you and schedule a 20-minute interview with you. In the interview, we will ask you to further explain your answers to some questions in the survey.

Is there any way being in this study could be bad for me?

There are no foreseeable risks involved in participating in this study. The possible risks or discomforts of this study are minimal.

Confidentiality

We will make every effort to limit the collections and disclosure of your personal information. Your survey answers will be stored in the Qualtrics cloud and will only be accessible to researchers of this study. We do not collect identifying information such as names or IP addresses. At the end of the survey, you will be asked if you are interested in participating in an additional interview via Zoom. If you choose to provide your email address, your survey responses may no longer be anonymous to the researcher. However, no names or identifying information would be included in any publications or presentations based on these data, and your responses to this survey will remain confidential.
Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team by contacting them at dbowman@vt.edu or feiyulu@vt.edu.

This research has been reviewed by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

• You have questions about your rights as a research subject
• Your questions, concerns, or complaints are not being answered by the research team
• You cannot reach the research team
• You want to talk to someone besides the research team to provide feedback about this research

By pressing “Continue” at the bottom right corner of this page, I indicate that:

☐ 1) I have read and understood the above information.
☐ 2) I consent to participating in this study.
☐ 3) I am at least 18 years of age.

By pressing “Continue” at the bottom right corner of this page, I indicate that:

☐ 1) I have read and understood the above information.
☐ 2) I consent to participating in this study.
☐ 3) I am at least 18 years of age.
Please fill out the background questionnaire

Gender

- Female
- Male
- Prefer not to say
- Other

Age

- 

Which country are you from?

- 

Where did you hear about this survey? (Please indicate the name of the Reddit community/Facebook group/mailing list if applicable)

- Reddit
- Facebook
- Twitter
What is your occupation? (if student, please indicate your level of study)

Please rate your level of experience with Augmented Reality:

Beginner 〇 〇 〇 〇 〇 Advanced

Which of the following devices do you own and use regularly?

☐ Smart phone
☐ Tablet
☐ Smart watch
☐ Personal computer
☐ Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
☐ Other

Watch the video prototype

Imagine that 10 years from now, you have a pair of lightweight,
powerful Augmented Reality (AR) glasses with all-day battery life. The glasses are designed to be an auxiliary tool to provide hands-free access to everyday information to assist your real-world activities. Below you will find a video of our envisioned interface for everyday use cases with such AR glasses. Please watch the video below. After you finish watching, click on the arrow to continue to the next section.

Video prototype of Glanceable AR Interfaces

Feedback on the video prototype

Please rate the system in the video you just viewed.

Strongly disagree  Somewhat disagree  Neither agree nor disagree  Somewhat agree  Strongly agree
| I think that I would like to use this system frequently | | | | | | |
|--------------------------------------------------------|---|---|---|---|---|
| The system seems unnecessarily complex.                | | | | | | |
| I think the system would be easy to use.               | | | | | | |
| I think that I would need the support of a technical person to be able to use this system. | | | | | | |
| I found the various functions in this system well integrated. | | | | | | |
| I think there was too much inconsistency in this system. | | | | | | |
| I would imagine that | | | | | | |
For each pair of words below, indicate your perception of the system in the video you just viewed.

most people would learn to use this system very quickly.

I would imagine the system being very cumbersome to use.

I would feel very confident using the system.

I would need to learn a lot of things before I could get going with this system.
Task-oriented questions

For the various tasks shown in the video, let's compare the ways you would currently do this (e.g., with tablets/mobile phones/smart watches) with the ways shown in the video using the AR glasses. The questions refer to six categories of tasks, defined below.

Please answer the following three questions about Task 1: Check information while stationary (e.g., check task list/check calendar/activity in front of a desktop)

What device would you currently use to complete Task 1?

☐ Smart phone
☐ Tablet
☐ Smart watch
Personal computer
☐ Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
☐ Other

How easy would it be to access information for Task 1, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th></th>
<th>1 (Very Hard)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Very Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of access - current way</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ease of access - in video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How much distraction from what you are doing would you feel when accessing information for Task 1, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th></th>
<th>1 (Very Distracting)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Not Distracting at all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction - current way</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distraction - in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Please answer the following **three questions** about **Task 2: Check information while moving (e.g., check weather/activity while exercising)**

What device would you currently use to complete **Task 2** when you do not have access to a personal computer?

- [ ] Smart phone
- [ ] Tablet
- [ ] Smart watch
- [ ] Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
- [ ] Other

How easy would it be to access information for **Task 2**, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th>Ease of access - current way</th>
<th>1 (Very Hard)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Very Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of access - in video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How much distraction from what you are doing would you feel when accessing information for Task 2, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th>1 (Very Distracting)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Not Distracting at all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction - current way</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
</tr>
<tr>
<td>Distraction - in video</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
<td>⬜</td>
</tr>
</tbody>
</table>

Please answer the following three questions about Task 3: Check information while hands are occupied (e.g., check recipe while cooking)

What device would you currently use to complete Task 3 when you do not have access to a personal computer?

☐ Smart phone
☐ Tablet
☐ Smart watch
☐ Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
☐ Other
How easy would it be to access information for **Task 3**, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th>1 (Very Hard)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Very Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of access - current way</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Ease of access - in video</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
</tbody>
</table>

How much distraction from what you are doing would you feel when accessing information for **Task 3**, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th>1 (Very Distracting)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Not Distracting at all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction - current way</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Distraction - in video</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
</tbody>
</table>

Please answer the following **three questions** about **Task 4: Interact**
with apps while stationary (e.g., set timer/mark task as completed/open link from email in front of a desktop)

What device would you currently use to complete Task 4?

☐ Smart phone
☐ Tablet
☐ Smart watch
☐ Personal computer
☐ Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
☐ Other

How easy would it be to interact in Task 4, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th></th>
<th>1 (Very Hard)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Very Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of interaction - current way</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Ease of interaction - in video</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
How much distraction from what you are doing would you feel when interacting for **Task 4**, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th>1 (Very Distracting)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Not Distracting at all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction - current way</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distraction - in video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please answer the following **three questions** about **Task 5: Interact with apps while moving (e.g., change songs/playlists while exercising)**

What device would you currently use to complete **Task 5** when you do not have access to a personal computer?

- [ ] Smart phone
- [ ] Tablet
- [ ] Smart watch
- [ ] Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
- [ ] Other
How easy would it be to interact in Task 5, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th></th>
<th>1 (Very Hard)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Very Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of interaction - current way</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Ease of interaction - in video</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

How much distraction from what you are doing would you feel when interacting for Task 5, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th></th>
<th>1 (Very Distracting)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Not Distracting at all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction - current way</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Distraction - in video</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Please answer the following three questions about Task 6: Interact
with apps while hands are occupied (e.g., set timer/flip recipe pages while cooking)

What device would you currently use to complete Task 6 when you do not have access to a personal computer?

☐ Smart phone
☐ Tablet
☐ Smart watch
☐ Virtual assistant (e.g., Siri, Amazon Alexa, Google Home)
☐ Other

How easy would it be to interact in Task 6, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th></th>
<th>1 (Very Hard)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Very Easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ease of interaction - current way</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Ease of interaction - in video</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
How much distraction from what you are doing would you feel when interacting for **Task 6**, both in the way you currently do it, and with the interface shown in the video?

<table>
<thead>
<tr>
<th>1 (Very Distracting)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7 (Not Distracting at all)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction - current way</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distraction - in video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Feature-oriented questions**

Please rate the following **Features (a-j)** of the interface shown in the video.

**Feature a: Glanceable apps dock at the physical edges of a physical monitor and can be accessed by glancing**
Glanceable AR apps can be docked at the edge of the physical monitor

For each pair of words below, indicate your perception of Feature a

Obstructive  ○ ○ ○ ○ ○  Supportive
Complicated  ○ ○ ○ ○ ○  Easy
Inefficient  ○ ○ ○ ○ ○  Efficient
Confusing  ○ ○ ○ ○ ○  Clear
Boring  ○ ○ ○ ○ ○  Exciting
Not interesting  ○ ○ ○ ○ ○  Interesting
Conventional  ○ ○ ○ ○ ○  Inventive
Usual  ○ ○ ○ ○ ○  Leading Edge

Feature b: Voice input/gestures control the apps being looked at by the user
For each pair of words below, indicate your perception of the **Feature b**

- Obstructive  
- Complicated  
- Inefficient  
- Confusing  
- Boring  
- Not interesting  
- Conventional  
- Usual  
- Supportive  
- Easy  
- Efficient  
- Clear  
- Exciting  
- Interesting  
- Inventive  
- Leading Edge

**Feature c: Notification appears at the central vision to guide the**
user to look at the app that has new information

For each pair of words below, indicate your perception of Feature c

Obstructive □ □ □ □ □ □       Supportive
Complicated □ □ □ □ □ □       Easy
Inefficient □ □ □ □ □ □       Efficient
Confusing □ □ □ □ □ □       Clear
Boring □ □ □ □ □ □       Exciting
Not interesting □ □ □ □ □ □       Interesting
Conventional □ □ □ □ □ □       Inventive
Usual □ □ □ □ □ □       Leading Edge
Feature d: Apps extends with more detail when being gazed at

For each pair of words below, indicate your perception of Feature d

| Obstructive      | Supportive       | Complicated    | Easy            | Inefficient | Efficient | Confusing | Clear     | Boring         | Exciting   | Not interesting | Interesting | Conventional | Inventive | Usual       | Leading Edge |
|------------------|------------------|----------------|-----------------|-------------|-----------|-----------|-----------|----------------|------------|----------------|-------------|--------------|-----------|------------|-------------|--------------|
**Feature e: Apps can be attached to the user for mobile use**

For each pair of words below, indicate your perception of **Feature e**

<table>
<thead>
<tr>
<th>Obstructive</th>
<th>Supportive</th>
<th>Complicated</th>
<th>Easy</th>
<th>Inefficient</th>
<th>Efficient</th>
<th>Confusing</th>
<th>Clear</th>
<th>Boring</th>
<th>Exciting</th>
<th>Not interesting</th>
<th>Interesting</th>
<th>Conventional</th>
<th>Inventive</th>
<th>Usual</th>
<th>Leading Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Feature f: While user is mobile, apps will stay at the periphery to be unobtrusive, but can be accessed through a quick glance

For each pair of words below, indicate your perception of Feature f

<table>
<thead>
<tr>
<th>Obstructive</th>
<th>Supportive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complicated</td>
<td>Easy</td>
</tr>
<tr>
<td>Inefficient</td>
<td>Efficient</td>
</tr>
<tr>
<td>Confusing</td>
<td>Clear</td>
</tr>
<tr>
<td>Boring</td>
<td>Exciting</td>
</tr>
<tr>
<td>Not interesting</td>
<td>Interesting</td>
</tr>
<tr>
<td>Conventional</td>
<td>Inventive</td>
</tr>
<tr>
<td>Usual</td>
<td>Leading Edge</td>
</tr>
</tbody>
</table>
Feature g: Apps will dock back to their original positions after the space is recognized

For each pair of words below, indicate your perception of **Feature g**

<table>
<thead>
<tr>
<th>Obstructive</th>
<th>☐ ☐ ☐ ☐ ☐ ☐</th>
<th>Supportive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complicated</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Easy</td>
</tr>
<tr>
<td>Inefficient</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Efficient</td>
</tr>
<tr>
<td>Confusing</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Clear</td>
</tr>
<tr>
<td>Boring</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Exciting</td>
</tr>
<tr>
<td>Not interesting</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Interesting</td>
</tr>
<tr>
<td>Conventional</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Inventive</td>
</tr>
<tr>
<td>Usual</td>
<td>☐ ☐ ☐ ☐ ☐ ☐</td>
<td>Leading Edge</td>
</tr>
</tbody>
</table>
Feature h: Share content from desktop to AR display to create a new Glanceable app

For each pair of words below, indicate your perception of Feature h

Obstructive ○ ○ ○ ○ ○ ○ Supportive
Complicated ○ ○ ○ ○ ○ ○ Easy
Inefficient ○ ○ ○ ○ ○ ○ Efficient
Confusing ○ ○ ○ ○ ○ ○ Clear
Boring ○ ○ ○ ○ ○ ○ Exciting
Not interesting ○ ○ ○ ○ ○ ○ Interesting
Conventional ○ ○ ○ ○ ○ ○ Inventive
Usual ○ ○ ○ ○ ○ ○ Leading Edge
**Feature i: Hands-free access to information while hands are occupied**

For each pair of words below, indicate your perception of **Feature i**

<table>
<thead>
<tr>
<th>Obstructive</th>
<th>Supportive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complicated</td>
<td>Easy</td>
</tr>
<tr>
<td>Inefficient</td>
<td>Efficient</td>
</tr>
<tr>
<td>Confusing</td>
<td>Clear</td>
</tr>
<tr>
<td>Boring</td>
<td>Exciting</td>
</tr>
<tr>
<td>Not interesting</td>
<td>Interesting</td>
</tr>
<tr>
<td>Conventional</td>
<td>Inventive</td>
</tr>
<tr>
<td>Usual</td>
<td>Leading Edge</td>
</tr>
</tbody>
</table>
Feature j: Open glanceable apps on desktop for more complex tasks

For each pair of words below, indicate your perception of Feature j

Obstructive  Supportive
Complicated  Easy
Inefficient  Efficient
Confusing  Clear
Boring  Exciting
Not interesting  Interesting
Conventional  Inventive
Usual  Leading Edge
Comments

Do you think Augmented Reality displays will become a major way to acquire everyday information in the future?

- Yes
- No
- Not sure

If lightweight, stylish AR glasses with the system shown in the video actually existed, would you like to use it in your daily life?

- Yes
- No
- Not sure

What would you like about having an AR interface like this? In what ways does the AR interface from the video seem better than what you have now?
What would you dislike about having an AR interface like this? In what ways does the AR interface from the video seem worse than what you have now?

If you were to use this AR interface in everyday life, are there other functions (not shown in the video) that you think should be included?

For what other scenarios (not shown in the video) do you think it would also be beneficial to use an AR system like this?

What would you change about the design of this AR interface?
Are there any issues with this AR interface that would make it difficult to implement, or difficult to use in other scenarios not shown in the video?

If you have any other comments about this AR interface or any of your answers above, please enter them here.

If you would like to be contacted for a follow-up interview via Zoom (20 minutes), please leave your email address below. We will contact you and schedule a time with you shortly.

We are devoted to gathering feedback from people with different
backgrounds and ages. Please feel free to share this survey to your family and friends! Thank you for your participation!
D.1.2 Interviews
Semi-structured interview questions:

1. Could you explain in more detail your answer to [Question X] in the survey?

2. Could you explain what do you mean by the [phrases] you mentioned in your answer to [Question X]?

3. If this kind of interface already exists, what feature do you perceive to be dominant to your frequent usages of such interface?

4. Do you have any further comments or suggestions you would like to make to the AR interface shown in the video?
D.1.3 IRB Approval Letter
MEMORANDUM

DATE: July 7, 2020

TO: Douglas Andrew Bowman, Feiyu Lu

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires October 29, 2024)

PROTOCOL TITLE: Glanceable AR: Evaluating User Perceptions on Future Augmented Reality Interfaces

IRB NUMBER: 20-519

Effective July 7, 2020, the Virginia Tech Human Research Protection Program (HRPP) determined that this protocol meets the criteria for exemption from IRB review under 45 CFR 46.104(d) category (ies) 2(ii).

Ongoing IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities impact the exempt determination, please submit an amendment to the HRPP for a determination.

This exempt determination does not apply to any collaborating institution(s). The Virginia Tech HRPP and IRB cannot provide an exemption that overrides the jurisdiction of a local IRB or other institutional mechanism for determining exemptions.

All investigators (listed above) are required to comply with the researcher requirements outlined at: https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Determined As: Exempt, under 45 CFR 46.104(d) category(ies) 2(ii)
Protocol Determination Date: July 7, 2020

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.
<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.
D.1.4 Video Demonstrations

The video prototype is available both online here and in the supplementary material files that are attached to this dissertation.

D.2 Capstone Implementation I & II

D.2.1 Informed Consent Form
Title of research study: IRB #20-652: Evaluating Augmented Reality Interfaces for Everyday Authentic Usages

Principal Investigator: Dr. Doug A. Bowman  Contact: (540) 231-2058 or dbowman@vt.edu

Other study contact(s): Feiyu Lu  Contact: (540) 257-4562 or feiyulu@vt.edu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form.

This research project is intended to evaluate the Augmented Reality system we implemented for everyday use. We will explore how users tend to use augmented reality to assist their everyday tasks, and understand users’ perceptions of future everyday augmented reality uses. In this study, participants will go through an initial zoom meeting with the experimenter, use an AR application we developed in their everyday scenarios for at least 2.5 hours, and finally complete an online survey and an interview. The AR application will display information such as your Google calendar events, incoming emails, Google fitness data and to-do lists to you. We will collect your feedback about how you feel, what you like/dislike about using AR to acquire general information. This research will help us understand user perceptions on future usages of augmented reality displays in different everyday contexts. The total time of this study will be 5 hours over the course of 5 days. Participants will be compensated for $75 after all study sessions are finished.

Detailed Information: The following is more detailed information about this study in addition to the information listed above.

Inclusion Criteria: The study population for this study includes people who: 1) are at least 18 years old, 2) are English speakers, 3) have uncorrected normal vision, 4) regularly use Google’s email and calendar services, and 5) have access to a Magic Leap One AR headset (if located outside the Blacksburg, VA area).

Who can I talk to?
If you have questions, concerns, or complaints, or think the research has hurt you, please send an email to Feiyu Lu at feiyulu@vt.edu or talk to the research team directly at Room 160 (Sandbox), Moss Arts Center, 190 Alumni Mall, Blacksburg, VA 24060

This research has been reviewed and approved by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject
- Your questions, concerns, or complaints are not being answered by the research team
- You cannot reach the research team
- You want to talk to someone besides the research team to provide feedback about this research

How many people will be studied?
We plan to include about 15 people in this research study.
Consent to Take Part in a Research Study

What should I know about being in a research study?
- Someone will explain this research study to you
- Whether or not you take part is up to you
- You can choose not to take part
- You can agree to take part and later change your mind
- Your decision will not be held against you
- You can ask all the questions you want before you decide

What happens if I say yes, I want to be in this research?
You will become a participant in this research. You will be invited to try out the AR system we developed, and give us your feedback about ways to interact with data and information in AR. The study will take place remotely; you will use the AR headset in your own home.

What are the experimental procedures?
Participants who are not physically present in the Blacksburg area will need to have access to the Magic Leap One AR headset to be able to participate in this study. For participants who are physically present in the Blacksburg area, we will provide a sanitized headset for them to use. After participants express interest in participating, we will provide the consent information to them via email in pdf format. For participants who are physically present in the Blacksburg area, the COVID-19 consent addendum will be provided as well. Participants will consent to participate by digitally signing the form(s) and email it back to the investigator. If participants consent that they are willing to participate in this study, for participants who are physically present in Blacksburg, we will first schedule a time for them to pick up the headset on campus. For participants who are not physically present in Blacksburg, they will use the headset they have access to participate. Next, we will hold a remote zoom meeting with the participant. In the remote zoom meeting, we will introduce the study goals and procedures, ask for the email address that participants use frequently every day to deploy the application, instruct participants to install the application on the Magic Leap One AR headset, and ask them to use the app for at least 2.5 hours total over the course of the next three days. Participants will be asked to fill out a Qualtrics diary survey to document their use and their perceptions about the app immediately after each usage session. After all sessions are covered, we will ask participants to submit their diaries, complete an online survey, and participate in a final zoom interview. Finally, for Blacksburg participants, we will schedule a time to collect the headset on campus. The total duration of this study will be around 5 hours across 5 days.

What is the total duration of this study?
The duration of an individual subject's participation will be 5 hours over the course of approximately one week.

Day 1: The first zoom call that will instruct participants remotely on how to use the hardware, how to install and use the AR system, and will provide instructions for study participation, will take around one and a half hour.

Days 2-4: Then participants will be asked to use the AR system for at least 2.5 hours total over the course of the next three days and keep a diary to document their uses.
Day 4 or 5: The online survey and final interview will take around 1 hour to finish.

Compensation

Participant will be compensated for $75 via Amazon gift card after all data collection procedures in the study have been finished.

What happens if I say yes, but I change my mind later?

You can leave the research at any time, for any reason, and it will not be held against you. All the data collected to the point of withdrawal will be properly discarded and destroyed immediately. Compensation will be pro-rated based on the number of hours you participate if you leave the study early. Compensation ranges from $15 to $75 based on the extent of participation. Participants who attend the initial zoom meeting will be compensated for $15. Then participants will receive $10 for completing every session of using the AR system (5 sessions in total). Participants who complete the entire study will receive a total amount of $75 compensation.

Is there any way being in this study could be bad for me? (Detailed Risks)

Using AR technology could produce symptoms of sickness or discomfort in some users. These symptoms are usually mild, and may include dizziness, nausea, eye strain, headache, or disorientation.

To enable personalized google services in AR, we will need you to grant access to the AR application to access your Google calendar, Gmail, Google Fit and Google Sheet content. The deployment of the app will be conducted via zoom and we will share our screen to ensure everything is transparent. In order to deploy the app, we may be able to briefly see one of your emails/calendar event, but after we finish deploying, we will remove the access token so we will not be able to access any of your google information. You will be able to disable access of the AR application to your google services at any time in your google account setting by going to “Manage your google account – Security – Manage third party access”, and remove the app named “ReadGoogleCalendar”.

We assure you that although the AR application will be accessing your Google account, this is only for the purpose of displaying email, calendar, google fitness and task information to you while using the app, and that the investigators will not be able to access your Google account or password at any time.

What happens to the information collected for the research?

We will make every effort to limit the use and disclosure of your personal information, including survey responses, diaries, and interview notes, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

We will need to collect your Google email address so that we can set up the AR application to access your Google email, calendar and fitness accounts. We will not ask for your password, nor will the research team ever have access to your email, calendar and fitness information. After the study is complete, we will remove your email address from our records.
Consent to Take Part in a Research Study

If identifiers are removed from your private information or samples that are collected during this research, that information or those samples could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

**Can I be removed from the research without my OK?**
Participants will not be removed from the study without their consent.

**What else do I need to know?**
This research is being funded by *Office of Naval Research via University of California, Santa Barbara*

We will not offer to share your individual test results with you. You may accept or decline these results.

Your information and samples (both identifiable and de-identified) might be used to create products or to deliver services, including some that may be sold and/or make money for others. If this happens, there are no plans to tell you, or to pay you, or to give any compensation to you or your family.

There are no significant benefits that individual subjects might experience from participating in this research. However, the subjects will get to experience novel Augmented Reality technology through the Magic Leap One headset.

**Signature Block for Capable Adult**

Your signature documents that you have read and understand the information outlined in this document and your permission to take part in this research. Please keep a digital signed copy of this form for your records.

<table>
<thead>
<tr>
<th>Signature of subject</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Printed name of subject</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature of person obtaining consent</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Printed name of person obtaining consent</td>
<td></td>
</tr>
</tbody>
</table>

Virginia Tech Institutional Review Board Protocol No. 20-652
Approved October 16, 2020
D.2.2 Questionnaires

Demographic Questionnaire
Block 1

Participants ID assigned to you (Ask the experimenter if unclear):

Gender

- Male
- Female
- Prefer not to say
- Other

Age

What is your occupation? (if student, please indicate your level of study)
Please rate your level of experience with Augmented Reality:

- Beginner
- Advanced

Please select which of the following Augmented Reality head-worn displays you have used:

- Microsoft HoloLens 1
- Microsoft HoloLens 2
- Magic Leap 1
- Google Glass
- Other

What types of devices do you use regularly (mark all that apply)?

- Smart phone
- Tablet
- Smart watch
- Personal computer
- Virtual assistant (e.g., Amazon Alexa, Google Home, Siri)
- Other
Do you think Augmented Reality displays will become a major way to acquire everyday information in the future?

- Yes
- No
- Not Sure
Diary Entry
Default Question Block

Participants ID (Ask the experimenter if unclear):

Period of use (Example: 2020/10/1 14:00-15:00)

On a scale of 1-10 (1 the worst and 10 the best), How would you rate your user experience in this session?

1 2 3 4 5 6 7 8 9 10

User Experience

Please try to answer the following questions based on your last session of use. It is not required to answer every question, whether to answer each question or not depends on your actual user experience. Please
put N/A for questions that you think are not applicable.

Where were you and what were you doing while using the app?

How often did you use the app to get information? What widgets did you use most often?

Where and how did you arrange the widgets?

Did you find the app useful during this session? Why or why not?

Did you find the app distracting during this session? Why or why not?
Describe one interesting occurrence during this session that illustrates the potential advantages of the Glanceable AR approach compared to devices you normally use (e.g., phones/smart watches)

Describe one interesting occurrence during this session that illustrates the potential downsides of the Glanceable AR approach compared to devices you normally use (e.g., phones/smart watches)

Overall, how do you feel about your use in this session?

Other thoughts (e.g., interesting things/issues happened during uses)
Post-study Questionnaire
Default Question Block

Participant ID assigned to you (Ask the experimenter if unclear):

Based on your experiences of using this AR prototype, please give ratings for the following statements. Assume that the system would run on a lightweight, comfortable, eyeglass-form-factor headset (i.e., ignore the current form of the AR hardware).

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>
I think that I would like to use this system frequently. | ○ | ○ | ○ | ○ | ○ |
I found the system | ○ | ○ | ○ | ○ | ○ |
unnecessarily complex.

I thought the system was easy to use.

I think that I would need the support of a technical person to be able to use this system.

I found the various functions in this system were well integrated.

I thought there was too much inconsistency in this system.

I would imagine that most people would learn to use this system very quickly.

I found the system very
Based on your experiences of using this AR prototype, for each pair of words below, indicate your perception of the system. Assume that the system would run on a lightweight, comfortable, eyeglass-form-factor headset (i.e., ignore the current form of the AR hardware).

- Cumbersome to use.
- I felt very confident using the system.
- I needed to learn a lot of things before I could get going with this system.

Annoying  ○ ○ ○ ○ ○ ○  Enjoyable
Not understandable  ○ ○ ○ ○ ○ ○  Understandable
Creative  ○ ○ ○ ○ ○ ○  Dull
Easy to learn  ○ ○ ○ ○ ○ ○  Difficult to learn
Valuable  ○ ○ ○ ○ ○ ○  Inferior
Boring  ○ ○ ○ ○ ○ ○  Exciting
Not interesting  ○ ○ ○ ○ ○ ○  Interesting
<table>
<thead>
<tr>
<th>Positive Trait</th>
<th>Score</th>
<th>Negative Trait</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unpredictable</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Predictable</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Fast</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Slow</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Inventive</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Conventional</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Obstructive</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Supportive</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Good</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Bad</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Complicated</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Easy</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Unlikable</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Pleasing</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Usual</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Leading Edge</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Unpleasant</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Pleasant</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Secure</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Not secure</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Motivating</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Demotivating</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Meets expectations</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Does not meet expectations</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Inefficient</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Efficient</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Clear</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Confusing</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Impractical</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Practical</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Organized</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Cluttered</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Attractive</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Unattractive</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
<tr>
<td>Friendly</td>
<td>☐ ☐ ☐ ☐ ☐</td>
<td>Unfriendly</td>
<td>☐ ☐ ☐ ☐ ☐</td>
</tr>
</tbody>
</table>
Conservative  ○ ○ ○ ○ ○ ○ Innovative

Powered by Qualtrics
Semi-structured Interview Scripts
Semi-structured interview questions:

1. Tell me about the different scenarios (situations, locations, and tasks) where you tried using the AR interface. For each scenario, what devices do you typically use to obtain information?

2. How do you feel about using the AR interface in each of these scenarios? What pros and cons do you perceive when using the AR interface compared to devices you frequently use?

3. Do you feel that the AR interface decreased use of your other devices? When you chose to use the AR interface instead of another device, how was using the AR interface better or worse?

4. Tell me about a specific situation where you found the AR interface to be very useful. Tell me about a specific situation where you found the AR interface to be unhelpful, distracting, or annoying.

5. How do you like the stationary and the follow modes in the interface? Are there any concerns you would have if you bring the AR apps with you?

6. How do you like the different interaction methods (gaze, gesture & voice) in the interface? Do you prefer different interaction methods in different scenarios?

7. Ignoring the form factor of the display, were there any issues you experienced in these scenarios while using the AR interface?

8. Could you explain what do you mean by the [phrases] you mentioned in your diary entry [X]?

9. In scenario [X], could you elaborate more on why you felt that the AR interface was [Y]?

10. Could you elaborate on what happened in your comment in diary entry [X]?

11. Are there other scenarios that you did not experience but you think you would like to use this AR interface in?
12. What changes would you make to the design of this AR interface? What other features would you wish to be included?

13. [Only in Capstone II] How did you feel about the interactivity features in the system, such as being able to browse pages, trash/star emails?

14. [Only in Capstone II] We implemented both Blink and Dwell to interact with the AR apps. Which one did you like more and why?

15. [Only in Capstone II] Were there particular strategies you used to arrange the AR apps?

16. [Only in Capstone II] How did you feel about using the system while having social interactions with others?

17. If the display was as light-weight and comfortable as a normal pair of eyeglasses, would you want to use the AR interface in your everyday life?

18. Do you have any further comments or suggestions you would like to make about the AR interface?
D.2.3 IRB Approval Letter
MEMORANDUM

DATE: October 16, 2020

TO: Douglas Andrew Bowman, Feiyu Lu

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires October 29, 2024)

PROTOCOL TITLE: Evaluating Augmented Reality Interfaces for Everyday Authentic Usages

IRB NUMBER: 20-652

Effective October 16, 2020, the Virginia Tech Institutional Review Board (IRB) approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at: https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Approved As: Expedited, under 45 CFR 46.110 category(ies) 4,6,7
Protocol Approval Date: October 16, 2020
Progress Review Date: October 16, 2021

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.
<table>
<thead>
<tr>
<th>Date*</th>
<th>OSP Number</th>
<th>Sponsor</th>
<th>Grant Comparison Conducted?</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/07/2020</td>
<td>PET47EVN</td>
<td>University of California, Santa Barbara (Title: View management and user interface optimization for wide-area mobile augmented real)</td>
<td>Compared on 10/14/2020</td>
</tr>
</tbody>
</table>

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.
D.2.4 Video Demonstrations

Capstone I

Video demonstrations of this research is available both online here and in the supplementary material files that are attached to this dissertation.

Capstone II

Video demonstrations of this research is available both online here and in the supplementary material files that are attached to this dissertation.