Rapid Design & Prototyping Methods for Mobile Head-Worn Mixed Reality (MR) Interface & Interaction Systems

Brady E. Redfearn

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
Industrial and Systems Engineering

Joseph L. Gabbard, Chair
C. Patrick Koelling
Deborah E. Dickerson
D. Scott McCrickard

January 2, 2017

Blacksburg, Virginia

Keywords: rapid prototyping, augmented reality (AR), virtual reality (VR), mixed reality (MR), user-centered design (UCD), User Experience (UX), human-computer interaction (HCI), head-worn display (HWD), design, interface, interaction

Rapid Design & Prototyping Methods for Mobile Head-Worn Mixed Reality (MR) Interface & Interaction Systems

Brady E. Redfearn

Academic Abstract

As Mixed Reality (MR) technologies become more prevalent, it is important for researchers to design and prototype the kinds of user interface and user interactions that are most effective for end-user consumers. Creating these standards now will aid in technology development and adoption in MR overall. In the current climate of this domain, however, the interface elements and user interaction styles are unique to each hardware and software vendor and are generally proprietary in nature. This results in confusion for consumers.

To explore the MR interface and interaction space, this research employed a series of standard user-centered design (UCD) methods to rapidly prototype 3D head-worn display (HWD) systems in the first responder domain. These methods were performed across a series of 13 experiments, resulting in an in-depth analysis of the most effective methods experienced herein and providing suggested paths forward for future researchers in 3D MR HWD systems.

Lessons learned from each individual method and across all of the experiments are shared. Several characteristics are defined and described as they relate to each experiment, including interface, interaction, and cost.

Rapid Design & Prototyping Methods for Mobile Head-Worn Mixed Reality (MR) Interface & Interaction Systems

Brady E. Redfearn

General Audience Abstract

Trends in technology development have shown that the inclusion of virtualized objects and worlds will become more popular in both professional workflows and personal entertainment. As these synthetic objects become easier to build and deploy in consumer devices, it will become increasingly important for a set of standard information elements (e.g., the "save" operation disk icon in desktop software) and user interaction motifs (e.g., "pinch and zoom" on touch screen interfaces) to be deployed in these types of futuristic technologies.

This research effort explores a series of rapid design and prototype methods that inform how a selection of common interface elements in the first responder domain should be communicated to the user. It also explores how users in this domain prefer to interact with futuristic technology systems. The results from this study are analyzed across a series of characteristics and suggestions are made on the most effective methods and experiments that should be used by future researchers in this domain.

Table of Contents

List of Figures

List of Tables

List of abbreviations

For clarification purposes, the following abbreviations are used throughout this research for common terms related to the various segments of this effort:

- Objectives 1-3
    - O1, O2, O3
- Research Questions 1-6
    - RQ1, RQ2,…, RQ6
- Experiences 1-13
    - E1, E2,…, E13
- Activities 1-10
    - A1, A2,…, A10

1: Introduction

1.1: Problem Statement

Augmented/Virtual/Mixed Reality (AR/VR/MR) three-dimensional (3D) systems have become increasingly more prevalent over the last several decades. Technological trends indicate they will continue to increase in popularity in the future due largely to the fact that previously expensive and complex research-based head-worn displays (HWDs) have begun to move to the consumer development space (Chi, Kang, & Wang, 2013; Zhou, Duh, & Billinghurst, 2008). This is a direct result of decreasing costs in 3D system hardware and software development. While many tools exist to rapidly design and prototype 2D interfaces in both desktop and mobile-based platforms (e.g., PowerPoint, Axure, Balsamiq, WireFlow), no such tools exist for 3D MR interfaces. This results in a much higher investment in the time and resources required to create a 3D interface when compared to a 2D system. And because the currently available MR systems are novel technologies that largely exist inside the academic/private industry research and development space, no standardized interface elements exist for these systems; hardware and software vendors are building proprietary elements. Additionally, no standardized interaction techniques exist for an MR system. It is necessary to prepare for the future increased use of these 3D interface systems by engaging in meaningful research efforts today that help enable user adoption of these technologies through the analysis of the methods and tools that currently exist and providing a post hoc evaluation of those approaches that are well-suited to rapid design and prototyping workflows.

Before useful 3D interface design tools can be built, standardized interface elements and interaction motifs need to be designed and tested. MR is also plagued by basic human factors challenges (e.g., divided attention, cognitive overload, occlusion of virtual/real world objects) that must be addressed throughout 3D interface development processes. End-users of MR interface systems are also challenged by limited methods for information exchange and collaboration across organizational members. The effect of rapid prototype MR interface development and interaction methods in a real-world applied performance setting is not yet fully understood. (R. Azuma et al., 2001)

1.2: Motivation

Mixed Reality (MR) is one of the most exciting novel technologies of the computer age. When properly designed and implemented, it can be used in myriad situations as a method of enhancing human performance and increasing our capacity to efficiently perform various tasks. When improperly designed and implemented, 3D interface systems have a negative effect on human performance. Research in applicable domains such as maintenance/repair (S. Henderson & Feiner, 2011), medical (Fuchs et al., 1998), communication (Kato & Billinghurst, 1999), and national defense (S. J. Henderson & Feiner, 2009) have been of great interest to academic, government, and private institutions. By overlaying 2D and 3D digital content in the physical reality, MR interfaces can provide increased situational awareness, especially in high-risk scenarios. One of these high-risk domains of interest to the research team is that of first responders (e.g., police, paramedics). However, the effect of this novel technology on human performance and experience is not yet fully comprehended. Novel MR technologies require a new approach to rapid prototyping for interface design and evaluation. (R. Azuma et al., 2001)

1.2.1: First Responder Domain

The title of “first responder” is an extremely broad label. In fact, the National First Responders Organization (NFRO), which is a professional group that represents this population of men and women, simply defines the title as “…any individual who runs toward an event rather than away” (NFRO, 2014). They include as part of this definition a non-exhaustive register of 90 job titles that constitute those who fall into this category at local, regional, and national levels; ranging from lifeguards and doctors, to soldiers and coroners. This domain is of interest to the research team because of the unique challenges that exist within the work context of first responders, in addition to the previously stated challenges of the 3D MR interface system domain. Some of the characteristics that apply to these types of positions include:

- high-risk, critical job duties where life-and-death decisions are often made (e.g., the first responder is meant to protect and serve the community)
- high stress levels effect the responder (e.g., due to the high-risk situations)
- little control over direct assignments (e.g., other people assign them their duties)

- individuals with whom they interact are often difficult to deal with (e.g., citizens are under duress)
- high-risk scenarios are dealt with daily (e.g., consistently and frequently)
- the world in which they work is often very negative (e.g., many citizens do not seek out their assistance)
- they are passionate and dedicated to their work (e.g., they genuinely want to improve their communities) (Kohan & O'Connor, 2002; W. Gary Howard, Heather Howard Donofrio, & James S. Boles, 2004, p. 381)

Additional complexity in first responder organizations arises when regional and national scenarios are also considered; a place where systems engineering principles must be considered. As a result of this complex high-risk environment, the safety of the first responder is of vital concern to the systems that support them. Many first responders (e.g., police, National Guard, military) are often working as a physically distributed group as well. When a specific scenario requires an intervention, team members often co-locate at an incident location and exchange information with each other. At present, radio systems, smartphones, and laptops allow for this exchange (Baber, Sharples, Boardman, Price, & Haniff, 2001; Jan Willem Streefkerk, Wiering, van Esch-Bussemakers, & Neerincx, 2008).

It is often the case that first responder scenarios have frequently changing goals and evolving information needs as emergency situations are dealt with. Because of the physical distribution of team members during their daily work tasks, overall situational awareness is challenging to maintain even when emergencies are "routine" in their nature (Baber et al., 2001; Ferscha, 2000; J. W. Streefkerk, 2011). The high-stress environment – in terms of time constraints and multiple tasks to perform in a single scenario – also results in a reduced amount of attention that can be dedicated to interacting with technological tools to support a specific response scenario (McCrickard, Catrambone, Chewar, & Stasko, 2003; McFarlane, 2002). It is also the case that cognitive characteristics, such as memory and subject-matter expertise, can influence how technological tools are utilized (Carroll, 1993). These myriad constraints require that future technological systems designed for first responders show a measurable improvement in both safety and situational awareness in order for end-users to see them as a benefit to their daily operations and adopt them as a part of their workflow. (Bailey, Konstan, & Carlis, 2000; Iqbal & Horvitz, 2007)

This work will explore the use of different methods in the research, design, and prototyping of mobile head-worn MR interfaces for first responders. Researchers will perform an analysis of the lessons learned from these various methods (e.g., semi-structured interviews, card sorting, Gesture Elicitation) in order to gather data on what approaches prove useful to developing MR interfaces. Specific recommendations that address the use of rapid prototyping methods for mobile head-worn MR interfaces will be made.

1.3: Research Purpose, Objectives, and Questions

The ultimate purpose of this study is to ***recommend rapid prototyping methods for mobile head-worn MR interfaces,*** using first responders (e.g., police, paramedics) as the user group of focus. For this purpose, this work aims to address three overall objectives:

- Identify ***information requirements*** for first responder MR systems
- Design and prototype MR ***user interfaces*** for first responders
- Design and prototype MR ***user interactions*** for first responders

Along the course of interface development, this work will also examine human factors research questions relevant to each objective. The following section enumerates the three main objectives (e.g., O1, O2, O3) of the study in more detail by stating the research questions (RQs) of the proposed study:

Objective 1: Information Requirements

- RQ1: What information do first responders expect to be available to them in a futuristic mobile HWD MR interface?
- RQ2: What information is most critical to the first responder to allow them to perform their job safely?

Objective 2: User Interfaces

- RQ3: How can critical interface information be communicated to first responders utilizing multiple modalities (e.g., visual, aural, haptic) of notification?
- RQ4: How does the context in which a user experiences a prototype effect the interface feedback they provide?

Objective 3: User Interactions

- RQ5: How do first responders desire to interact with critical information?
- RQ6: How does the context in which a user experiences a prototype effect the interaction feedback they provide?

1.4: Significance of the Study

First, this research will provide the guidance necessary to make the appropriate tradeoffs – in terms of scope, schedule, and resources – when designing and prototyping MR interfaces that will enable future researchers to rapidly design and prototype their own 3D MR experiences, such as is commonly performed in 2D interface design and prototyping today. Second, this research effort will provide guidance on specific information element representations that could be used in the first responder domain. Third, this research will provide specific guidance on interaction motifs that could be used in this domain.

A post hoc analysis from the information gathering, design, and prototyping methods of the MR experiences employed herein will ultimately lead to a set of recommended rapid prototyping methods for MR interfaces, which do not currently exist. The specific characteristics of these experiences will be detailed in order to enable the reader to select any of the studied experience types and make their own determination as to which one suits the needs of their own research agenda.

1.5: Method Overview

Rapid prototyping methods for mobile head-worn MR interfaces present unique challenges for researchers and practitioners alike. While these MR systems are becoming more ubiquitous, it is still not fully understand how the implications of these systems influence user groups. Moreover, as this work looks to the future, existing rapid prototyping methods for 2D interface systems may not adequately address the needs of 3D interface systems.

This research effort will analyze a series of rapid design and prototyping iterations in the development of a mobile head-worn MR interface for first responders. Each of these iterations will rely on existing user-centered design (UCD) and rapid prototyping methods. Each sub-section of the overall research effort will include a careful post-iteration breakdown and analysis that includes the lessons learned from each implementation of each method and what worked well and/or did not work well during each phase. This collection of takeaways will then feed into a series of recommended rapid prototyping method characteristics for mobile head-worn MR interfaces that take the most beneficial tools and methods used throughout this research effort and provide a way forward for other 3D interface rapid prototype research efforts to build upon. The overall progress of the effort can be communicated with the following Figure, which displays a progression of the three research Objectives (i.e., O1, O2, O3) discussed in detail throughout this paper:

Figure 1: Research Effort Overview

1.5.1: O1: Information Requirements

This effort will first follow a basic information-gathering process that will inform all of the design and prototype iterations going forward. Through the use of semi-structured interviews with subject matter experts (SMEs), researchers can better understand the domain of application for first responders (e.g., police and paramedics). This understanding will help to create several research artifacts to inform future work.

With a general understanding of the research domain landscape and how the user currently performs their duties, an enumeration of information elements can be performed through the use of User Experience Design (UxD)-based semi-structured interviews and Participatory Design (PD)-based storytelling methods to gather the information elements that are currently on-hand to the SME. With a greater understanding of the current state of information elements, the research team can then brainstorm with SMEs and collect a “wish list” of information elements that could be perceived as useful to the SMEs in their future professional duties.

Lastly, semi-structured interviews and storytelling methods are coupled with a card sorting method. This is utilized by the research team in order to categorize and prioritize the list of possible information elements in order to focus the research effort on what elements are perceived as most useful to SMEs during a common response scenario in a rapid prototyping method for mobile head-worn MR interfaces. The research artifacts of this Objective will be used to answer RQ1 and RQ2. A post-mortem analysis of the methods and tools utilized during O1 will then be shared in this paper.

1.5.2: O2: User Interfaces

With the completion of the initial information gathering process in O1 that addresses the RQs of that Objective, the research team will perform a series of design and prototype iterations during which user interfaces for mobile head-worn MR systems will be explored. Each iteration will progressively include a larger body of work in both design and prototype methods as the effort progresses. The final user interface research artifacts of this Objective will address RQ3 and RQ4. A post-mortem analysis of the methods and tools utilized during each user interface design and prototype iteration of this research will follow.

1.5.3: O3: User Interactions

With the completion of O2, the research team will perform a final series of iterative design and prototype iterations, which address user interactions for mobile head-worn MR systems. Each design and prototype user interaction iteration will consider all of the previous work performed to be essential to each future iteration. As before, a post-mortem analysis of the methods and tools utilized during each user interaction rapid prototyping iteration will be performed and reported in this research paper.

1.5.4: Discussion and Recommendations

Finally, a post hoc analysis and discussion of the entire research effort will be performed. A recommended set of design and prototype methods for 3D MR HWDs will be presented to the reader. These methods will be informed by the analyses performed throughout this research effort and will allow the reader to select the best experience type that meets their own research objectives. This series of recommendations should help future researchers to avoid the pitfalls and shortcomings of specific experiential characteristics that have been felt throughout this longitudinal research effort.

2: Literature Review

Previous research in areas of interest to this effort illustrate both the existing problems of rapid design and prototyping methods for 3D MR interface systems and some theoretical approaches that are applicable to aiding in and experimenting with methods in this domain. The following background research and development subject matter topics have been applied throughout this effort:

2.1: Development of Longitudinal User-Centered Design (UCD)

The earliest identified report of a long-term usability study in computer software was published over twenty years ago (Cook, Science, & Science, 1994), but the subject has gained more popularity in the twenty-first century with various articles and conference papers discussing the topic. In 2002, for example, the results of a six-week evaluation of Microsoft Word were published, sharing a comparison of multiple interface designs that were tested and evaluated in a realistic field study to gather feedback on the usability of Word (McGrenere, Baecker, & Booth, 2002). Several years later, a study of scientific databases was published that stressed the importance of having a usable system to “provide a basis for the assessment of data quality and [the] possibility of data sharing between scientists” (Hueni, Nieke, Schopfer, Kneubuhler, & Itten, 2009, p. 565). The Journal of Engineering Design published a 28-day study on the usability of a home appliance, with participants ranging 30-82 years old, that attempted to describe a methodology for tracking long-term user data by way of user diaries (Imai et al., 2010). These and other studies describe a variety of long-term UCD definitions and research methods, but none of these methods have been consistently applied across the studies found in the existing body of academic research, nor were there theoretical foundations found underlying these methodologies to allow the research efforts to be easily comparable among each other.

In 2006, Hornbaek published an article in the International Journal of Human-Computer Studies (IJHCS) that analyzed the current status of usability research as defined by currently published works from core HCI forums and remarked that “[t]he studies reviewed show that users typically interact only briefly with interfaces under investigation…” (Hornbæk, 2006, p. 93). In fact, of the 180 studies analyzed in that paper, only 13 of them lasted more than five hours (Hornbæk, 2006). In 2010, the necessity for long-term usability was stressed to the academic community once again explaining that further research is critical "because we are approaching the 'loyalty decade', where interaction experience will become the main success factor [for organizations] (Jakob Nielsen, 2008)" (Alghamdi, 2010; Law & van Schaik, 2010, p. 313; Jakob Nielsen, 2008; Van Schaik & Ling, 2011). The need for "a clear picture of how UX changes over time …[is necessary given that]… user-expectation and user-affect dynamically evolve with the actual usage of the product over time" (Law & van Schaik, 2010, p. 314). Searching through the citations of this article showed that only three academic journal articles have even mentioned this longitudinal call for research in the seven years since it was published (Clarivate Analytics, 2017).

The cited literature, in summary, shows that long-term UCD is an important field of interest to both academic and practitioner's areas of research and practice, but most UCD studies still only capture brief periods of time (minutes and hours) and are not longitudinal in their duration. The reasons for this apparent lack in long-term UCD research are outside of the scope of this research study, but it is clear there is room for the exploration of the effects of long-term UCD practices in a real-world rapid design and prototyping method for mobile head-worn MR interfaces scenario.

2.2: Usability Methodologies

Because long-term UCD is a new field of study, the research literature indicates that more longitudinal methodologies are needed to test, refine, and produce more user-centric systems. Because there is a lack of long-term UCD methodologies and practices that are in popular practice in the UCD domain, an alternative approach must be found. One method of addressing the longitudinal UCD process is to begin with existing short-term testing methodologies and expand on them to incorporate long-term attributes. While there are many applications of short-term testing methodologies, there is very little research in longitudinal UCD and related methodologies for use in applied settings. The reasons for a lack of long-term UCD testing methods is outside the scope of this report (as was mentioned in the previous section), but the lack of many well established, tested, and proven long-term UCD methodologies indicate that short-term approaches are currently more popular in research and practice. It has been shown that most published research utilizes traditional, short-term, cross-sectional UCD testing scenarios. A rational approach to develop a long-term UCD methodology is to build upon existing short-term practices as a foundation to research and modify them for long-term use. This approach will be particularly effective if the chosen short-term methods lend themselves to continuous or cyclically repeated applications. In the latter case, it is useful if the last stage of one cycle flows easily into the first stage of the next cycle and if there are mechanisms to preserve and maintain long-term knowledge and trends. One strong candidate for short-term testing scenarios are those proposed by Tullis and Albert (2010). Tullis and Albert suggest that there are ten common usability study scenarios that can be used to collect and analyze usability data, as shown in the following Figure (Tullis & Albert, 2010, p. 50):

Table 1: Ten Common Usability Study Scenarios

Scenario Number	***Scenario Title***
1	Completing a transaction
2	Comparing products
3	Evaluating the frequent use of the same product
4	Evaluating navigation and/or information architecture
5	Increasing awareness
6	Problem discovery
7	Maximizing usability for a critical product
8	Creating an overall positive user experience
9	Evaluating the impact of subtle changes
10	Comparing alternative designs

These methodologies are popular to use in cross-sectional UCD testing scenarios, but little data was available for review on long-term UCD studies that included consideration for these ten methods of gathering user data (Hornbæk, 2006). Because these scenarios have been successfully applied in traditional short-term UCD testing (Tullis & Albert, 2010), they provided a foundational testing basis for this long-term UCD study.

Some of the testing scenarios in the previous Figure are more relevant to this study than others. In particular, Scenarios 1-8, and 10 were most important due to the nature of the scenarios for first responders and the practical application of the scenarios to long-term UCD testing principles in a rapid prototyping method for mobile head-worn MR interfaces.

2.3: Usability Engineering Lifecycle

Originally published by Deborah J. Mayhew (1999) in a text subtitled as “a practitioners handbook for user interface design", the Usability Engineering Lifecycle has been a common process to apply to UCD efforts around the world. First presented in its formalized usage to the CHCI conference in 1998 (Mayhew, 1998), this was a concept and phrase originally conceived and published by Jakob Nielsen six years earlier (1992), although Mayhew significantly expanded the idea by 1999. The following Figure shows this model in its 1999 form:

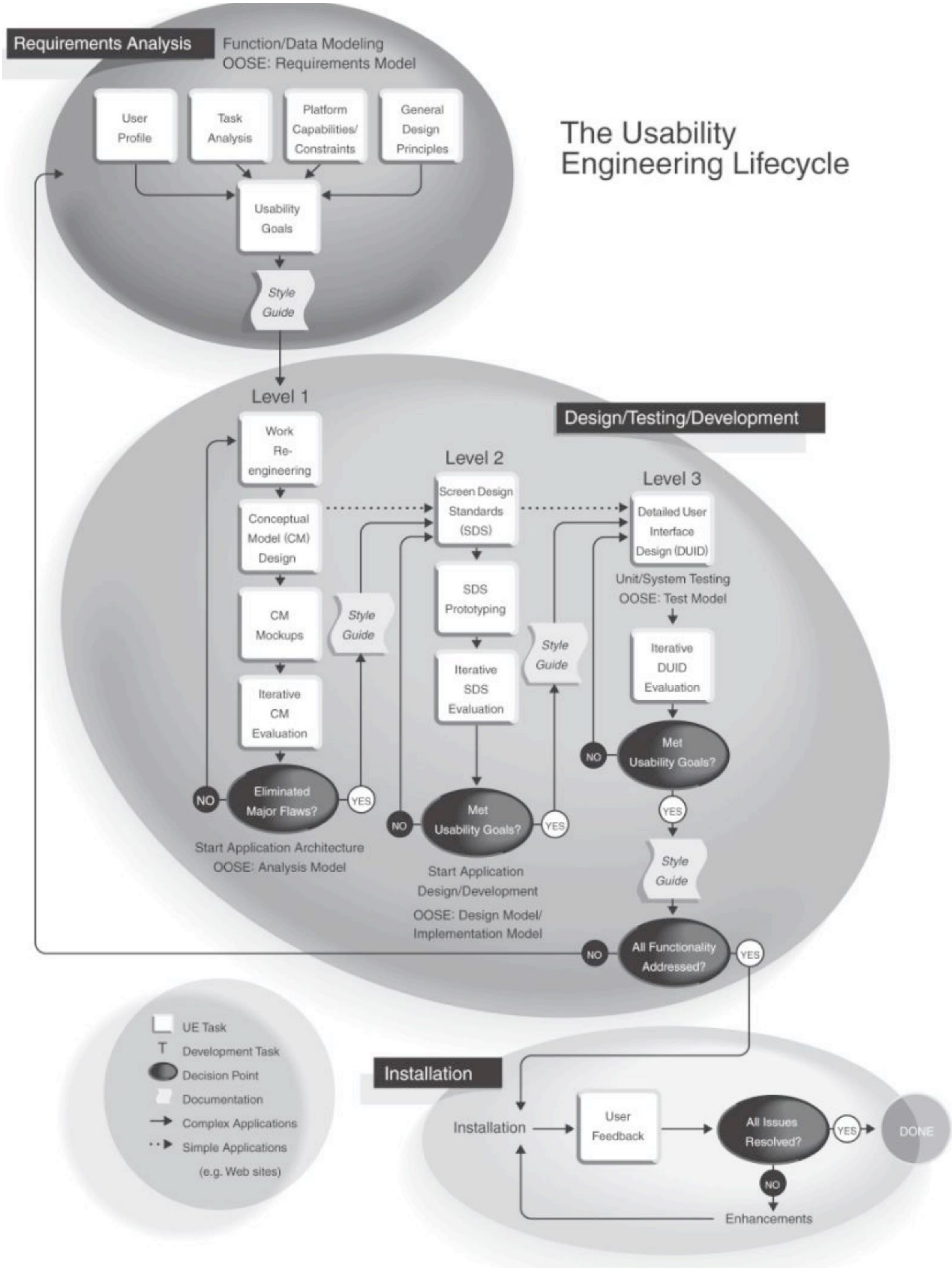


Figure 2: The Usability Engineering Lifecycle (Mayhew, 1999)

While this model has three main phases that describe the process (i.e., requirements analysis, design/testing development, installation), it is also composed of many micro-level tasks and processes that make a very comprehensive UCD methodology with great detail at each step as to what the practitioner needs to do to create a successful product. It also includes several iterative tasks and processes as well so that the practitioner can repeat many steps as needed until the effort is complete. However, this process includes a clear step labeled “done” at the end of the lifecycle. The research team argues that such a step indicates a short-term perspective to the lifecycle and while one could suggest that merely starting the lifecycle process over again can reengage the benefits of the lifecycle, this seems contradictory to the perspective of longitudinal ethnographic research studies where the constant change of the environment and people in it actually require a constant vigilance be maintained with the UCD system at all times; there is never truly a state of “done.” (Kaptelinin & Nardi, 2009; Mayhew, 1999)

Strictly longitudinal UCD studies, therefore, would seem to require a dramatic change in practice is needed; where an attitude of continuous improvement, much like processes found in successful manufacturing environments, would need to be analyzed as a template for how to integrate the perspective of uninterrupted UCD improvement into a business process that has been traditionally applied as a style of momentary and single-intervention enhancements for decades. (Tennant, 2001) The Usability Engineering Lifecycle might be helpful for a single effort, as a temporary consultant would have with a client, but lacks the longitudinal perspective, knowledge transfer, and continual process improvement parameters that are necessary to the future of UCD practices in order to secure customers in the “loyalty decade.” (Alghamdi, 2010; Law & van Schaik, 2010, p. 313; Jakob Nielsen, 2008; Van Schaik & Ling, 2011)

2.4: The Wheel

Another popular approach to modern UCD processes was published by Hartson and Pyla (2012) in *The UX Book*. This process is known as “The Wheel,” which is shown in the following Figure:

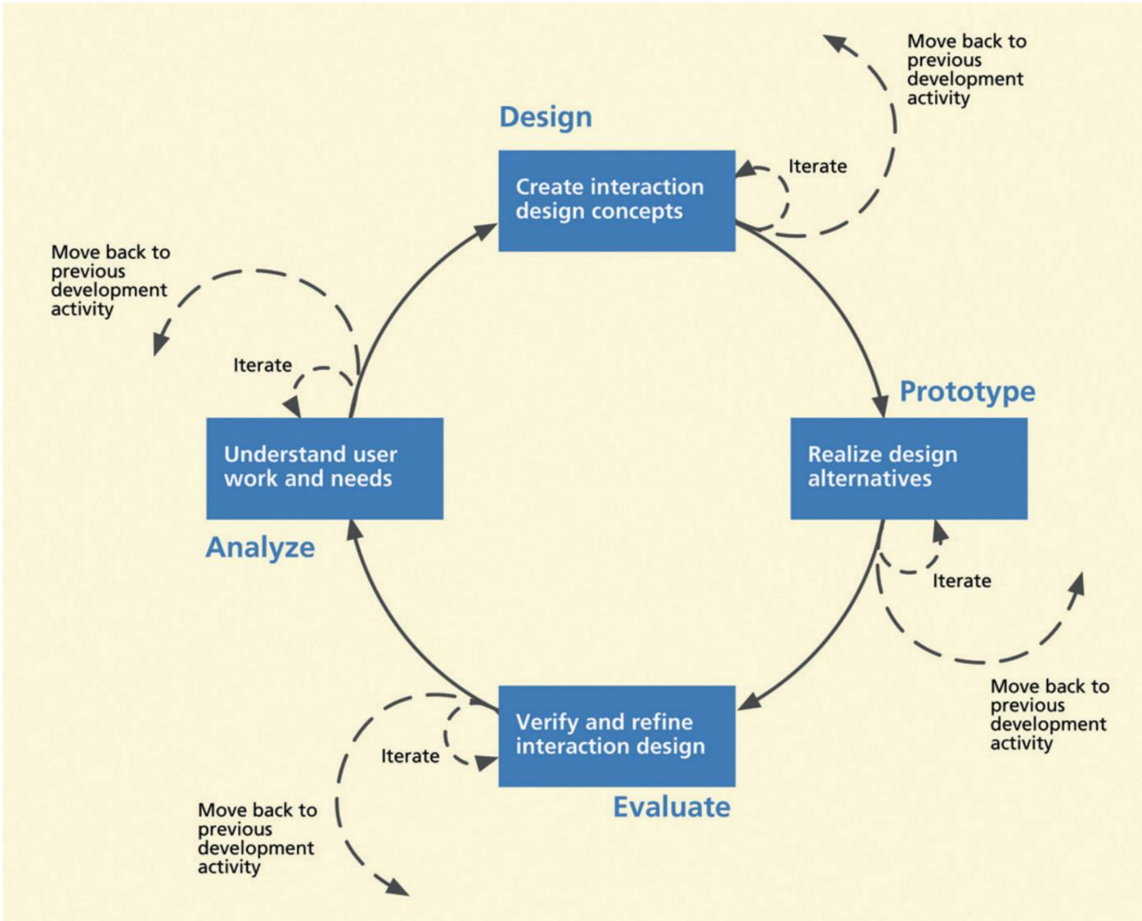


Figure 3: The Wheel, (Hartson & Pyla, 2012)

The Wheel provides a much different approach to UCD practice compared to the Usability Engineering Lifecycle. Instead of providing a well-structured and clearly defined set of processes and individual tasks, The Wheel offers a perspective of continuous improvement through a very simple process of four steps: design, prototype, evaluate, analyze. “The primary overall objective of The Wheel lifecycle process is to keep moving forward and eventually to complete the design process and make the transition to production” (Hartson & Pyla, 2012, p. 57). This perspective seems well-aligned to the longitudinal perspective needed in UCD design methods, however, The Wheel is still bound by a defined “transition to production.” In the modern technology development domain of Computer Science (which is where The Wheel originates), the research team suggests there is no need for such a formal transition. All manner of design and prototype changes, therefore, can be put in production at any time as part of a UCD process (i.e., Google's Gmail product was labeled "beta" for years) (Smith & Caruso, 2010). And multiple versions of such changes can be put in production simultaneously (e.g., parallel software development/incremental testing performed and deployed by Amazon, Google, and others) to test what designs work well for users in a live, non-laboratory, truly ethnographic setting (Myers, Sandler, & Badgett, 2011). Additionally, users can choose interface versions themselves or fully customize them to inform the software developers on what is preferred (e.g., revert to old system if you do not like the new one).

The Wheel may serve as a well-designed lifecycle system for those UCD practitioners who appreciate brevity, but it is still a process that relies on historical paradigms that often no longer apply to modern and future-forward technology development processes. It is based on traditionally slow software development processes. The research team suggests that while this model is a better fit for a longitudinal UCD development process, that it is still missing the ethnographic activity-based and knowledge management components that are important to long-term UCD processes that will become more important in the future as customer loyalty becomes an ever-increasingly prominent metric in technology development. (Kaptelinin & Nardi, 2009)

2.5: Case Studies in Research Design

Case studies have been defined by Gerring (2004, p. 341) as "…an intensive study of a single unit with an aim to generalize across a larger set of units." They were first used historically in the physical sciences and began to transition into practice in the social sciences in the early 1800s when Frederic Le Play began work on what he would eventually publish half a century later, *Les ouvriers européens* (1878). Case study methods have since been applied in myriad individual domains, including economics, political science, sociology, psychology, business, law, software development, HCI, Usability, User Experience, and related fields. However, case studies are not without their critics and some researchers argue they contribute very little to understanding the world around us (Achen & Snidal, 1989; Lieberson, 1991, 1994). (Gerring, 2004; Healy, 1947, pp. 97–98)

2.5.1: The History and Evolution of Case Study Research

As mentioned previously, case-based research methods have been practiced for over 200 years in modern science (Frédéric Le Play, 1878; Healy, 1947). With such a long tradition of practice, it should be no surprise that case studies seem to have permeated every scientific domain in practice today. Being an ambiguous term, a "case study" can refer to an effort that is qualitative in nature and of a small population sample; or research that is ethnographic, observational, and outside of a laboratory setting (Yin, 1994). Gerring (2004) adds the following additional categories of possible case study types: research that is characterized by process-tracing, an investigation of the properties of a single case, or an investigation of "…a single phenomenon, instance, or example," the last type of which Gerring believes to be "the most common usage" (2004, pp. 341–342). What this seems to indicate is that because there are so many applied uses of the term "case study" across myriad domains, then the definition of a case study and the methods used to describe them vary as widely as their application. This unstructured and indistinct phrase, therefore, results in research efforts that can be published and practiced anywhere on the scientific contribution spectrum of beneficial to injurious.

2.5.2: Limitations of Case Studies

While it is a very popular method in use today, many academics are quite critical of case studies and believe them to contribute very little to the progress of scientific research (Achen & Snidal, 1989). Part of this abhorrence of case studies seems to stem from the lack of standardization of a method (Gerring, 2004; Thomas, 2011). In fact, Gerring concludes that "practitioners continue to ply their trade but have difficulty articulating what it is that they are doing, methodologically speaking. The case study survives in a curious methodological limbo" (2004, p. 341). Researchers have begun to call for more structure and thought to be put into what a case study truly is to make it a more academically rigorous method altogether. (Gerring, 2004; Starman, 2013; Thomas, 2011)

Miles (1979) describes the type of qualitative data gathered from a case study "as an attractive nuisance." He further explains that his three largest complaints of qualitative data are as follows: (a) "the actual process of analysis during case-writing was essentially intuitive, primitive, and unmanageable" (Miles, 1979, p. 597), (b) "cases usually required a considerable amount of revision to take account of the factual errors, the defensive responses, and the genuinely alternative interpretations" (Miles, 1979, p. 597), and (c) "the art of cross-site analysis is even less well-formulated than within-site analysis" (Miles, 1979, p. 599). He concludes his 1979 paper by stating that without appropriately rigorous and structured scientific research methods, case studies serve as little more than "story-telling" (Miles, 1979, p. 600).

Flyvbjerg (2006) presents a slightly more structured description of the shortcomings of case-based research with a list of what he calls "five common misunderstandings." While a proponent of case-based research, he agrees with many authors that poorly executed case studies are what present the real danger to the scientific validity of such a research strategy. The biggest limitations of case studies in his view are as follows:

- Theoretical knowledge is emphasized as more valuable than practical knowledge
- Single-case studies are not generalizable
    - Only multiple-case studies yield scientific contributions
- Case studies are more useful for generating hypotheses
    - Other methods are better suited for testing hypotheses and building theories
- Case studies are highly biased
    - Authors want their views to be proven correct
- It's difficult to summarize a case study (Flyvbjerg, 2006, p. 219)

In other words, an improperly structured and executed case study is a very popular research output in the published literature. These low-quality academic studies provide an endless stream of fodder for critics of case study research. Following in the same vein as Miles, Flyvbjerg's critiques and conclusions are also quite valid. However, what Miles fails to address are any possible solutions to resolve his initial concerns. A more structured and scientific analysis phase, along with more rigorous data gathering processes can serve to mitigate many of his original complaints with case-based research. In fact, Yin (1981) responds to Miles' rebuke of case study research two years later with an equally intrusive reprimand of his own, combined with several suggestions on how to appropriately structure a case study in order to improve the scientific contribution to society with case-based research.

2.5.3: Ideal Applications of Case Studies

While authors such as Miles have been very critical of case study methods, Yin and other proponents of case-based research provide responses to their peer's critiques. In fact, Yin (1981, p. 58) begins his response to Miles' earlier article by drawing the reader's attention to the fact that Miles' article "is an example of a frequent confusion regarding types of evidence (e.g., qualitative data), types of data collection methods (e.g., ethnography), and research strategies (e.g., case studies)." This distinction of three separate definitions has not been generally adopted across case study research even today. In fact, Yin is stating that there is no such thing as a *case study method*; it is a *case study research strategy* by its very definition. The exact classification of whether a case study is a method vs. a research strategy in the social sciences is currently of less import to the research team's progress than gathering together the best suggestions and practices across several social science domains in order to determine which case study applications are most ideal to future research efforts. However, further exploration of this distinction should prove useful in ensuring the research team provide a beneficial contribution to the academic literature. These subtle differences should be clearly defined during this effort. (Yin, 1981)

Flyvbjerg (2006) answers his five critiques of low-quality case-based research (as described earlier) with the following abbreviated retorts:

- Concrete, context-dependent knowledge *is* valuable.

- One *can* generalize on the basis of a single case.
- The case study *is* useful for generating and testing of hypotheses.
- The case study contains *no greater bias* than other methods of inquiry.
- The problems in summarizing case studies are due more often to the *properties of the reality studied* than to the case study as a research method. (emphasis added) (Flyvbjerg, 2006)

In other words, there is no reason why a well-documented and structured case study that follows established methodologies published and reviewed in the scientific literature is any less effective or beneficial to describe a research effort. In fact, the case study fills an important niche in scientific experimentation and has been shown time and again to be "…a method that holds up well when compared to other methods in the gamut of social science research methodology" (Flyvbjerg, 2006, p. 241).

Case studies are a method well-suited for understanding a particular set of relationships, surrounded by a deep understanding of a set of participants in a specific environment. Case studies are unique and versatile in their approach due to the scope of the method; it is at the discretion of the author to decide whether to focus on a single person as the unit of study, an entire organization, or anything in between. And the measures and techniques used in this approach are varied and unique to the study at hand. However, it is clear that some case studies have had such an impact that they are cited thousands of times as important works of research (Eisenhardt, 1989; Flyvbjerg, 2006; Stake, 2009; Yin, 1981, 1994). By incorporating the practices of historically influential case studies into this effort, the research team can increase the validity of its research process and the positive impact of the research output on the academic literature.

"…The case study is probably best understood as an ideal-type rather than a method with hard-and-fast rules" (Gerring, 2004, p. 346). Gerring argues that there are important methodological ambiguities that cannot be eradicated from a case study: the type of inference (descriptive vs. causal), the scope of proposition (depth vs. breath), the unit of homogeneity (is the chosen population generalizable?), causal relationships (if there are any), the strategy of research (exploratory vs. confirmatory), the useful variance (single vs. many units), and the relevant ontologies (assumptions about how the world actually works) (2004). Thomas (2011) provides a visual of how seven different authors (or groups of authors) define the different kinds of case studies in research, as shown in the following Figure:

George and Bennett (2005, drawing on Eckstein, 1975)	Merriam (1988)	Stake (1995)	Bassey (1999)	de Vaus (2001)	Mitchell (2006) (drawing on Eckstein, 1975)	Yin (2009)
Theory testing	Descriptive	Intrinsic	Theory seeking	Descriptive/ explanatory	Illustrative	Critical
Atheoretical/ configurative-idiographic	Interpretative	Instrumental	Theory testing	Theory testing/ theory building	Social analytic	Extreme/ unique
Disciplined configurative	Evaluative	Single/ collective	Storytelling	Single/multiple case	Extended (over time)	Longitudinal
Heuristic	—	—	Picture drawing	Holistic/embedded	Configurative-idiographic	Representative
Plausibility probes	—	—	Evaluative	Parallel/sequential	Disciplined-configurative	Revelatory
"Building block" studies	—	—	—	Retrospective/ prospective	Heuristic	—
					Plausibility probes	

Figure 4: Types of Case Studies (Thomas, 2011, p. 516)

With so many different category names, despite many of them to be overlapping or with somewhat similar descriptions, it is clear that researchers have not reached a common narrative to describe what a case study is. Addressing these ambiguities and providing clear definitions will be essential to explore and address as this research effort progresses in order to ensure an ideal application of a case study method. "In all, the state of the art [with regards to case-based research] is not as impoverished as one might at first think, and case study practice can be dramatically improved by applying what is already known" (Yin, 1981, p. 64).

2.5.4: Established Methodologies for Case Studies

The research team has been slowly building a personalized methodology for this case study through the literature review process. An important outcome of this writing exercise is to begin to build a theory to test, complete with the selection of research questions and the formal hypotheses that will be tested in the research effort. Eisenhart (1989) provides an eight-step process for building a theory from case study research that will certainly prove valuable to the research team, as shown in the following Figure:

Step	Activity	Reason
Getting Started	Definition of research question	Focuses efforts
	Possibly a priori constructs	Provides better grounding of construct measures
	Neither theory nor hypotheses	Retains theoretical flexibility
Selecting Cases	Specified population	Constrains extraneous variation and sharpens external validity
	Theoretical, not random, sampling	Focuses efforts on theoretically useful cases—i.e., those that replicate or extend theory by filling conceptual categories
Crafting Instruments and Protocols	Multiple data collection methods	Strengthens grounding of theory by triangulation of evidence
	Qualitative and quantitative data combined	Synergistic view of evidence
	Multiple investigators	Fosters divergent perspectives and strengthens grounding
Entering the Field	Overlap data collection and analysis, including field notes	Speeds analyses and reveals helpful adjustments to data collection
	Flexible and opportunistic data collection methods	Allows investigators to take advantage of emergent themes and unique case features
Analyzing Data	Within-case analysis	Gains familiarity with data and preliminary theory generation
	Cross-case pattern search using divergent techniques	Forces investigators to look beyond initial impressions and see evidence thru multiple lenses
Shaping Hypotheses	Iterative tabulation of evidence for each construct	Sharpens construct definition, validity, and measurability
	Replication, not sampling, logic across cases	Confirms, extends, and sharpens theory
	Search evidence for "why" behind relationships	Builds internal validity
Enfolding Literature	Comparison with conflicting literature	Builds internal validity, raises theoretical level, and sharpens construct definitions
	Comparison with similar literature	Sharpens generalizability, improves construct definition, and raises theoretical level
Reaching Closure	Theoretical saturation when possible	Ends process when marginal improvement becomes small

Figure 5: Building Theory from Cast Study Research (Eisenhardt, 1989, p. 533)

This theory-building method should be a productive guide for the research team in formulating the beginning stages of the research effort, starting with the definition of research questions and concluding with testable hypotheses that can be executed in an experimental setting. Additionally, these testable theories will be tied to empirical evidence and because of this development process, Eisenhardt suggests her process "…is particularly well-suited to new research areas or research areas for which existing theory seems inadequate" (Eisenhardt, 1989, pp. 548–549)

Keeping in mind that Yin (1981) would describe a *case-based research strategy* as something apart from a unique *data collection method* (e.g., ethnography), further research is needed to define established methods that have been recognized as most useful for case studies. He describes several necessary characteristics of high-quality case-based research as follows:

- "Case studies can be done by using either qualitative or quantitative evidence" (Yin, 1981, p. 58). The research team plans to use both types of evidence to increase the validity of the study.
- "…The case study [does not] imply the use of a particular data collection method" (Yin, 1981, p. 59). Researchers plan to use several data collection methods to compare user responses across multiple studies.
- Case studies represent research strategies. The research team plans to apply its methods in real-life contexts, as explained through the use of Activity Theory, especially as "..the boundaries between phenomenon and context are not clearly evident" (Yin, 1981, p. 59).

Thomas (2011, p. 513) suggests "…that a case study must comprise two elements: (1) a 'practical, historical unity,' which I shall call the *subject* of the case study, and (2) an analytical or theoretical frame, which I shall call the *object* of the study." Much like Gerring (2004), Thomas provides the following definition for a case study:

Case studies are analyses of persons, events, decisions, periods, projects, policies, institutions, or other systems that are studied holistically by one or more methods. The case that is the subject of the inquiry will be an instance of a class of phenomena that provides an analytical frame—an object—within which the study is conducted and which the case illuminates and explicates. (Thomas, 2011, p. 513)

Defining the individual units of study, along with the external environment that influences it, is essential to a well-designed case study. Using Thomas' vernacular, a *subject* of this research study could be a police SWAT team and the *object* could be the process by which SWAT team member training effects their perception of a futuristic technology. Another important characteristic of an established case study method is to remember that the object can change throughout a study, but "…it is important to have some notion of a potential object in mind when the study begins and not to confuse it with the subject" (Thomas, 2011, p. 514).

Although a case study is often very ambiguous at the start of the effort, Thomas warns that "open endedness is [often] extended to an unwarranted expectation of structural looseness" (2011, p. 519). In other words, because a case study can be open-ended in its structure and method, that does not mean that there is a complete absence of structure altogether. A plan must be made ahead of the research effort and established methods must be executed (at least to some degree) or the case study will be little more than a nice-sounding story. (Gerring, 2004; Thomas, 2011)

2.5.5: How to Report Case Studies

The importance of case studies cannot be underestimated. Although some authors do not approve of them as a valid scientific method or research strategy, the research team believes that a highly-structured approach and conformance to published methodical processes are what set a good case study apart from the bad. Proponents of case-based research reiterate that the importance of publishing a "…systematic production of exemplars…" is essential to a thriving discipline and that the lack thereof is indication that a domain has little to offer practitioners in the real world (Flyvbjerg, 2006, p. 242).

By following well-established theory-building methods (Eisenhardt, 1989), the research team can appropriately judge the quality of published case studies that can serve as architypes for this effort. "Strong studies are those which present interesting or framebreaking theories which meet the tests of good theory or concept development (e.g., parsimony, testability, logical coherence) and are grounded in convincing evidence" (Eisenhardt, 1989, p. 549). By using high-quality examples of well-reported case studies, the research team can determine which portions of the data gathered during a case study are most relevant to the reporting of the experiment and will aid in helping the reader understand the scientific method executed in the effort.

Most case study reports are long-form stories that seem to involve no formal structure in how they are described. "This pitfall may be avoided if a study is built on a clear conceptual framework" (Yin, 1981, p. 64). Yin also suggests that the story-telling habits of case-based research could "…be replaced by a series of answers to a set of open-ended questions…" (1981, p. 64). However, Flyvbjerg (2006) argues that a long narrative in the case study can often show that a specific problem in the system is merely complex. Enough detail must be provided that the reader can understand the situation well without exhausting the reader with superfluous information that does not impact the subject and object of study. Peattie (2001) warns against summarizing case study research: "It is simply that the very value of the case study, the contextual and interpenetrating nature of forces, is lost when one tries to sum up in large and mutually exclusive concepts" (2001, p. 260). (Thomas, 2011)

2.6: Case Studies in UCD

The study of cases has been in practice for decades in the Information Systems (IS) domain. Benbasat et al. (1987) described the fit of case-based research to the IS domain as "particularly appropriate for certain types of problems: those in which research and theory are at their early, formative stages [44], and 'sticky, practice-based problems where the experiences of the actors are important and the context of action is critical' [4]" (Benbasat et al., 1987, p. 369). Much like the IS domain, UCD "is characterized by constant technological change and innovation….[UCD] researchers, therefore, often find themselves trailing behind practitioners in proposing changes or in evaluating methods for developing new systems" (Benbasat et al., 1987, p. 370). The value that case studies provide is summarized by these same authors into three main attributes:

- performing studies in a natural setting and developing theory from practice
- researchers can explore the process in depth
- when few previous studies have been carried out, this is a good method to apply (Benbasat et al., 1987)

While it is common practice today to perform a case study in many social science domains, the UCD community appears to be somewhat resistant to wide adoption of this practice. This, however, "presents an opportunity. Information Systems uses these approaches widely," (Kjeldskov & Graham, 2003, p. 326) and a single proven case study method has not yet been found that is obvious to apply to this specific effort. This is likely due to the "sticky, practice-based problems" (Benbasat et al., 1987, p. 369) often explored in a case study, along with other ambiguous-sounding adjectives used by other authors when describing case-based research studies and methods (Berg, 1998; Budwig, Jeong, & Kelkar, 2009; Kim et al., 2008; Stake, 1995; Yin, 1984). Stake (1995) even refers to performing case study research as an art form. In fact, many modern texts (Baxter, Courage, & Caine, 2015; Hartson & Pyla, 2012; Karapanos, Jain, & Hassenzahl, 2012; Kujala, Roto, Väänänen-Vainio-Mattila, Karapanos, & Sinnelä, 2011; Ryan & Potts, 2015; Tullis & Albert, 2010) discuss many facets of performing an actual case study (e.g., how to gather data, which metrics to record, how to analyze the study), but fail to outline a specific methodology to follow during the research process. It appears many authors still rely on the basic case study methods described in the IS, Design Science (DS), psychology, business, and other domains, from whence these practices originated (Martin, Hanington, & Hanington, 2012, p. 28). Benbasat et al. (1987, p. 371) provides a long-form process to apply in a case study that includes 11 characteristics, a much briefer method description for performing case studies specifically in the area of UCD is given by Martin et al. (2012):

- determine a problem

- make initial hypotheses
- conduct research through interviews, observations, and other forms of information gathering
- revise hypotheses and theory
- tell a story (Martin et al., 2012, p. 28)

This simple five-step process is simply an abbreviated version of the list given by Benbasat et al. (1987). It remains apparent that this very broad guidance will have to be thoroughly analyzed by the research team to create a successful rapid prototyping method for mobile head-worn MR interfaces for this longitudinal UCD effort.

2.7: Types of Error

While some UCD practitioners believe the issues that arise from usability are the same unit of study as "errors," Tullis and Albert describe usability issues to be "…the underlying cause of a problem, whereas one or more errors are a possible outcome" (Tullis & Albert, 2010, p. 81). In other words, error is the output of a process that is fed by poor usability inputs. Error rates, especially when combined with additional metrics, like task completion times, can provide a powerful method of conveying how many mistakes were made, where those mistakes were made in a process/workflow, how the design influences specific types of errors, and how usable a product really can be. (Tullis & Albert, 2010, p. 81)

Many types of error exist in the domain of systems design, including errors in measurement (Bland & Altman, 1996), inference to the general population (Bland & Altman, 2015), uncertainty and sampling (Altman & Bland, 2014a), precision and accuracy (Altman & Bland, 2014b), missing data points (Vickers & Altman, 2013), self-reporting biases (A. Adams, Soumerai, Lomas, & Ross-Degnan, 1999; Brener, Billy, & Grady, 2003; Donaldson & Grant-Vallone, 2002; Huizinga & Elliott, 1986), mathematical calculations (Altman & Bland, 2011), etc. Within the confines of traditional UCD, some errors become more prevalent than others, especially because a human is involved as part of the system. "Generally, an error is any action that prevents the user from completing a task in the most efficient manner" (Tullis & Albert, 2010, p. 81). These same authors explain that any deviation in the primary task completion or any inefficiency is therefore categorized as an error. If taken to an extreme, this could mean that when a subject is interrupted from completing a primary task (i.e., writing a research report) by an unrelated secondary task (i.e., the answering of a telephone call), then one or more errors have occurred. (Box, Hunter, & Hunter, 2005)

2.7.1: Effects of Error

While the user-based actions that cause an error to occur are varied and correlate directly to the task being measured, a longitudinal UCD approach has yet to identify whether or not specific types error will be increased or decreased throughout the effort because of the long-term nature of the study. Inferences could be made from high-quality and high-performing manufacturing and statistics best practices, like the Six Sigma process improvement system, that while a performance level of three sigmas (i.e., 99.99966%) can be expected in the short-term when a system is running efficiently using this specific improvement method, when a longitudinal perspective is analyzed, performance rates will degrade and an error rate of about 7% is to be expected in the long-term (Tennant, 2001, p. 26). This general concept of the degrading performance of a system over time is aligned with the concepts of ethnographic UCD research and Activity Theory-based principles as discussed previously in this research; due to the ever-evolving nature of humans and the tools they use.

Minimizing the effects of error, therefore, plays a key role in this research effort at all times. The following paragraphs provide examples on how the research team plans to reduce error when possible, including some practical examples of data gathering and analysis. The example list of some system error types in the previous section of this report will be used again here: measurement, inference to the general population, uncertainty and sampling, precision and accuracy, missing data points, self-reporting biases, mathematical calculations, etc. All methods of error reduction will come from high-quality peer-reviewed academic publications whenever possible, but one or more methods may need to be adapted to this unique domain of longitudinal UCD practice in a rapid prototyping environment for mobile head-worn MR interfaces.

2.7.2: Inference to the General Population

It is common practice in UCD to develop user profiles and user-based scenarios to help inform system design and keep product development focused on the end-user customer. “I am not the user” is a common mantra for practitioners to cite when making design decisions. Therefore, accurate representative user groups and descriptions are essential to reducing errors of inference to the general population. In this research effort, the rapid prototyping methods for mobile head-worn MR interfaces are not meant for the general population to use, but are targeted to a specific subset of users. The expert users, therefore, must play an integral part throughout the product development lifecycle in order to ensure that the end-user customer’s needs are met and inference errors do not creep into the effort. (Bland & Altman, 2015; Tullis & Albert, 2010, pp. 57–58)

2.7.3: Uncertainty and Sampling

Related to the previous section, it is essential that the research team has an accurate sampling of the users for whom the system will be designed. The users must also be fully integrated into the product development lifecycle in order to ensure that all questions are answered and user feedback is constantly and consistently gathered throughout the effort so there is consensus among the users that the proper system design has been created for their various tasks. (Altman & Bland, 2014a; Tullis & Albert, 2010, pp. 58–60)

2.7.4: Precision and Accuracy

All materials must be uniform (e.g., semi-structured interviews, surveys, verbal questions, visual design elements) in terms of how they are presented to the user. The same words and pictures must be given to every participant. Pilot studies will be utilized in order to finalize a testing script so that the experiment can be given precisely and accurately every time, to every user. (Altman & Bland, 2014b)

2.7.5: Missing Data Points

Vickers and Altman (2013) suggest that "the most straightforward approach is simply to ignore" a subject with missing data and perform a "complete case analysis," but this results in reduced statistical power, due to the decreased data sample size. There are several other methods that can be utilized as well, including last observation carried forward and multiple imputation, but "analysis of missing data teaches us the importance of avoiding missing data in the first place: an informed guess, even using a technique as sophisticated as multiple imputation, is still a guess" (Vickers & Altman, 2013). From these conclusions, diligent attempts at ensuring that data collection methods are successful and redundant, along with ensuring a sample size large enough that a small data loss has little effect, are the methods that will be employed by the research team to remove the errors that result in missing data points.

2.7.6: Self-reporting Biases

"Self-reported data give you the most important information about users' perception of the system and their interaction with it" (Tullis & Albert, 2010, p. 123). When the perception of a system is an important metric to gather, self-reporting methods will be very useful to this rapid prototyping method for mobile head-worn MR interfaces (e.g., "Do you like how the system visually looks to you?") (e.g., "Was the task easy to perform?"). When numerical measures that do not involve human perception are necessary, alternate metrics will be gathered (e.g., task completion time, error rates). (A. Adams et al., 1999; Brener et al., 2003; Donaldson & Grant-Vallone, 2002; Huizinga & Elliott, 1986)

2.7.7: Mathematical Calculations

It is important to rely on mathematical tools that are redundant and not error-prone in this research effort. When making calculations, it is important to first, gather the data in an accurate manner (i.e., the same measure is applied to all participants); second, record the data in an accurate manner (i.e., the measure is entered in a spreadsheet); third, accurate mathematical formulas are applied to the data (i.e., mean calculation instead of mode for the average age of users). To be certain calculations are performed properly, it is also good practice to check the data points a second time to ensure they are correct. (Altman & Bland, 2011)

2.7.8: Statistical Errors

Perhaps more helpful to the UCD practitioner is the list of nine steps given by Good and Hardin (2012, pp. 3–4) that one can perform in order to prevent errors in statistical work. These steps include a well-planned procedure that will be implemented in this research effort to immediately decrease statistical error, including:

- Set objectives and research intentions before conducting any work
- Define the population
- Recognize that what you are investigating may have chaotic components
- List all sources of variation and control them or measure them
- Formulate hypotheses and all of the associated alternatives
- Describe how to draw participants from the population
- Use estimators that are impartial, consistent, efficient, robust; with minimum loss
- Know the assumptions that underlie the tests you use
- Give the complete details of everything you did (Good & Hardin, 2012, pp. 4–5)

"Three concepts are fundamental to the design of experiments and surveys: variation, population, and sample" (Good & Hardin, 2012, p. 5). All of these steps will help to reduce and control the variation error, population error, and sample error that is expected to be present in human-based research and experimentation. As has been discussed in previous sections of this research effort, by thoughtfully applying the structured design processes practiced in the UCD domain (e.g., user profiles, user scenarios, storyboards) and making strategic error-avoidance process plans early, myriad categories of error can be dramatically reduced and/or eliminated throughout this research effort.

2.8: Gesture

While a very common word today, the etymology of the term "gesture" has its modern roots in late Middle English and was derived from the more ancient Latin root "gesta," meaning actions or exploits. Medieval Latin modified the word to "gestura" during the early 15th century. The related Latin word "gerere" added the terms "bear, wield, perform" to the meaning of "gestura" by the Middle Ages. While spelled with a "g" in the 1550s, the shortened version of "jest" was not meant as a joke, as it means today, but a "notable exploit" and was used as "a narrative of someone's deeds" ("Gesture: Definition of Gesture in Oxford Dictionary (American English) (US)," 2016). "The original sense [of the word gesture] was 'bearing, deportment', hence 'the use of posture and bodily movements for effect in oratory'" ("Gesture: Definition of Gesture in Oxford Dictionary (American English) (US)," 2016). While the term "gesture" might conjure images of a person who speaks dramatically by moving their hands around or the negative middle finger-raised response one might be given from cutting a car off in traffic, it can be more broadly described as any physical action used to express an idea. ("Gesture: Definition of Gesture in Oxford Dictionary (American English) (US)," 2016, "Online Etymology Dictionary," 2016)

Gestures will be a very important part of this research effort. Because this work includes an analysis of the actions of human beings in a longitudinal study that interacts with tools, technologies, and other people, it will be imperative to not only record the things people say, but also the things they do. This is not limited to general action-related categories (e.g., a subject presses a button), but less-obvious actions that convey important gestural meanings as well, such as a seemingly simple action performed that forces user discomfort (e.g., the subject presses a button after having to walk five steps away from their workspace first). Two physical actions that might have a similar outcome (i.e., a button press), can also have important impacts on user performance and/or satisfaction because the gestural processes are actually different in their real-world execution.

Gesture identification and execution become even more essential when one considers this as a possible input method to a technology system. Since the 1980s, research groups have performed studies that have consistently shown a "versatility and ease of use that can enter upon the management of graphic space with voice and gesture" (Bolt, 1980). This research effort will rely on semi-redundant, multi-modal input methods in order to allow users to perform tasks using well-designed systems that are both efficient and usable in a longitudinal manner, including a hybrid method of expert- and user-driven gesture identification schemes.

2.8.1: Gesture Identification

"Willingly or not, humans, when in co-presence, continuously inform one another about their intentions, interests, feelings and ideas by means of visible bodily action" (Kendon, 2004, p. 1). In fact, research shows that gesture is such an integral part of human communication that people who are congenitally blind use gestures, even with other people who are blind, because it is so tightly coupled to human interaction, despite there being no obvious benefit to such physical expressions (Iverson & Goldin-Meadow, 1998, p. 228). "Gestures therefore require neither a model nor an observant partner" (Iverson & Goldin-Meadow, 1998, p. 228). Because all humans use gestures to communicate, one might assume that gesture recognition systems are well-suited for controlling a technology system. However, the prevalence of gestures as a means of interaction has historically presented the opposite challenge: high levels of false-positive triggers. Consider the development of a gesture (e.g., eye blink) that triggers an action (e.g., the command "take a picture"). While the gesture might work for some users, those persons with more active eye movement or people who physically squint their eyes more often might trigger the action inadvertently. In fact, one study showed that while users were observed in such a gesture recognition scenario throughout five activity classes (reading a book, talking with another person, watching a video, solving mathematical equations, and sawing wood), that a recognition accuracy of 67% was observed (Ishimaru et al., 2014). When combined with a secondary gesture (i.e., a specific head motion pattern), the accuracy of the gesture recognition was 82% (Ishimaru et al., 2014). If a mission-critical gesture recognition rate of 99% or higher were needed for the command "take a picture" (e.g., for a first responder to record evidence at a crime scene), this type of reliability would simply be unacceptable to the user at even the 82% rate (Ishimaru et al., 2014).

Research and development of advanced algorithms and Artificial Intelligence (AI) training methods have been working to address the reliability of positive recognitions and the reduction of false-positive recognitions for decades, but some approaches are still error-prone (Dalal & Triggs, 2005). An appropriate level of detail must be given to the methods used to identify gestures in any modern technology system development, especially when performing rapid prototyping sessions with users (Vatavu, Anthony, & Wobbrock, 2012). By relying on proven modern scientific approaches such as "computer vision and pattern recognition techniques, involving feature extraction, object detection, clustering, and classification, [which] have been successfully used for many gesture recognition systems," (Mitra & Acharya, 2007, p. 312) along with "image-processing techniques such as analysis and detection of shape, texture, color, motion, optical flow, image enhancement, segmentation, and contour modeling, [which] have also been found to be effective" (Mitra & Acharya, 2007, p. 312), the research team can expect to develop an appropriate gesture identification system for this effort. (Dalal & Triggs, 2005; Mitra & Acharya, 2007)

Two general identification schemas could be used to define a given set of actions that are part of a gesture library available to the end-user customer; expert- and user-driven collections. The following sections of this research effort describe these two methods of identification, along with a proposed hybrid solution that is well-suited for this research effort.

2.8.1.1: Expert-Driven Identification

It is common practice in engineering efforts for an engineer building a system to be deemed (perhaps more likely self-proclaimed) the "expert" of that system. While it may be true that they understand how to build a product, they are not generally the end-user of the system. UCD design processes hold paramount the mantra, "I am not the user." This means that, above all, the actual people for whom a product is being developed must be involved throughout the entire product development lifecycle in order to ensure that the system is aligned with the real needs of the end-user. This also means that stakeholder experts from all relevant parts of an organization (e.g., first responder support personnel, supervisors, high-level management) should be in consistent and constant communication with the research team. By integrating the UCD experts and the sponsoring organization experts, following the UCD-based principles outlined previously with influence from ethnographic research and Activity Theory-based principles, expert-driven gesture libraries can be developed and standardized in the targeted user group. (Gibet, Courty, & Kamp, 2006; Kaptelinin & Nardi, 2009; Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007)

Reliance on expert-driven gesture identification systems and methods, such as those currently practiced in Augmented Reality (AR) ("A Survey of Augmented Reality Technologies, Applications and Limitations," 2010; R. Azuma et al., 2001; Kato & Billinghurst, 1999; Zhou et al., 2008), in team collaboration settings (Bragdon, DeLine, Hinckley, & Morris, 2011), or using touch screen systems (Findlater, Lee, & Wobbrock, 2012), are important to the success of this study. This is the traditional approach for identifying inputs. One such study cites an important expert-driven conclusion that states, "…that gestures are better suited for multi-tasking situations because they are less interruptive than touch interaction to users' primary tasks and are subjectively preferred by users in certain situations" (Karam, Lee, Rose, Quek, & McCrickard, 2009, p. 7). However, another study argues that expert-driven gesture systems might be less preferred by users, "…because professionals tend to generate more physically and conceptually complex gestures" (Morris et al., 2014, p. 42). The research team must perform a tradeoff analysis comparing several possible input modalities at every given moment of time throughout a process in order to determine whether a gesture is even appropriate for performing the task at hand or if a different modality is more fitting for the user. (Morris, Wobbrock, & Wilson, 2010)

2.8.1.2: User-Driven Identification

The importance of constant and consistent user-driven feedback has already been discussed throughout this research paper in order to ensure technology development is aligned with the real-world needs of the end-user customer. In terms of gesture identification, however, an emerging doctrine suggests that a technological system should be capable of being completely customized for every user of the system (Liu, Zhong, Wickramasuriya, & Vasudevan, 2009). This concept allows for the interaction library of one user (e.g., Bob) to be completely different from another user (e.g., Andy). Both systems will perform the same tasks and processes, but can utilize individual preferences to be precisely aligned with each end-user's personal preferences. If Bob prefers a gesture (e.g., waving his hand in the air) to perform a specific task (e.g., send an email), while Andy prefers a voice command (e.g., giving the verbal command "Send Message"), the user-driven identification system is aware of whom it is interacting with and can allow for both commands to be used for the same single task. (Kane, Wobbrock, & Ladner, 2011; Liu et al., 2009)

One method of extracting user-driven interactions that is growing in popularity is gesture elicitation (Morris et al., 2014, p. 41; Wobbrock, Morris, & Wilson, 2009). Perhaps the largest advantage of the gesture elicitation technique "…is that the technique is not limited to current sensing technologies; it enables interaction designers to focus on end users' desires as opposed to settling for what is technically convenient at the moment" (Morris et al., 2014, p. 42). Morris et al. also concludes that this method leads to "end-user involvement [that] can result in gesture sets that are more likely to be discoverable by and memorable to a large user base" (Morris et al., 2014, p. 42). The method also exhibits several pitfalls, especially a legacy technology bias (e.g., the user interactions in use today on a smartphone are the same interactions as the "futuristic" user responses), but these biases can be addressed through future research opportunities, including this research effort. (Morris et al., 2014; Wobbrock et al., 2009)

2.9: Augmented/Virtual/Mixed Reality (AR/VR/MR)

Virtual Reality (VR) is a technology that "…completely [immerses] a user inside a synthetic environment" (R. T. Azuma, 1997, p. 355). Part of this complete immersion includes the full occlusion of the real world so that only virtual objects appear to the user. In contrast, Augmented Reality (AR) is "…any system that has the following three characteristics:

- Combines real and virtual
- Is interactive in real time

- Is registered in three dimensions" (R. T. Azuma, 1997, p. 356)

This inclusion of real and virtual objects being combined together, along with see-through display of the real world set AR apart from VR systems. Since the late 90s, experts have argued that AR "…is far behind virtual environments in maturity" (R. T. Azuma, 1997, p. 379) due to the more difficult challenges that arise from combining real and virtual objects (e.g., augmentation method, resolution, field of view, object processing speed, focus, contrast, portability). However, "AR can potentially apply to all senses, including hearing, touch, and smell" (R. Azuma et al., 2001, p. 34), which potentially make an AR-based experience much more engaging and realistic to the user. During the 90s, most AR systems were focused only on the information that was displayed to a potential user and didn't "…significantly concern themselves with how potential users would interact with these systems. Prototypes that supported interaction often based their interfaces on desktop metaphors…or adapted design from virtual environments research…" (R. Azuma et al., 2001, p. 37). Little evaluation of interaction with AR systems has been done since Azuma's 2001 evaluation of the field. (R. Azuma et al., 2001; R. T. Azuma, 1997)

The need for further research in AR remains a necessity. "We need a better understanding of how to display data to a user and how the user should interact with the data" (R. Azuma et al., 2001, p. 43). Much of the more recent AR research has focused on perception and other low-level issues, but AR includes many high-level tasks as well, especially in the context of real-world UCD applications. In 2008, a review of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR) was published (Zhou et al., 2008). In the ten years preceding this review, including a total of 313 published papers, only 46 of them (<15%) addresses the topic of interaction techniques (Zhou et al., 2008, p. 194). This review also concluded two overarching issues with AR systems; "from a human factors point of view, there are also plenty of issues to be considered. Physically, the design of the system is often cumbersome…. Cognitively, the complex design of the system often makes it hard to use" (Zhou et al., 2008, pp. 198–199). Even at this time, some researchers were arguing that these two issues were not necessary to address until more technical problems were solved, but Zhou concludes "it should never be too early to consider…" physical and cognitive issues with AR systems (Zhou et al., 2008, p. 199) and adds that social, economic, and cultural issues must also be part of the design of AR systems. Subsequent literature surveys have identified similar needs and have provided similar conclusions in the last decade. (Feiten, Wolf, Oh, Seo, & Kim, 2005; Rabbi & Ullah, 2013; Sanna & Manuri, 2016; Soo Kyun Kim, Shin-Jin Kang, Yoo-Joo Choi, Min-Hyung Choi, & Min Hong, 2017; Van Krevelen & Poelman, 2010)

The phrase "Mixed Reality" (MR) serves as the most popular reincarnated reference to AR, although the title was first used by Milgram with the publication of the Virtuality Continuum (VC) and Reality-Virtuality Continuum (RVC) over two decades ago (1994; 1995). The VC describes an infinitely-variable scale between real-world and completely virtual environments along which a technology can potentially exist. In essence, MR is a more generic umbrella word to describe both VR- and AR-based systems. Because myriad technologies now exist that have some level of virtual fidelity, MR simply "…[involves] the merging of real and virtual worlds…" to any degree. This has resulted in a resurgence of MR as a taxonomy of choice to describe HWD, fixed-screen, and other augmented/virtual environment technologies. It will be used as the phrase of choice for this research effort because multiple technological systems that utilize varying degrees of realism and virtuality were utilized throughout this paper, depending on the scope, schedule, and resources available during the rapid prototype design and prototype phases. (Billinghurst, Clark, & Lee, 2015; Milgram & Kishino, 1994; Milgram et al., 1995; Van Krevelen & Poelman, 2010)

2.10: Participatory Design (PD)

Participatory Design (PD), unlike the many social science-based qualitative approaches used in UCD, was developed as a political and social movement in Scandinavia. As terms of a design approach, it "…often is viewed primarily as a set of methods and techniques for involving users in design" (Blomberg & Burrell, 2009, p. 88) and remaining committed to involving those users throughout the entire technology development lifecycle (Kensing & Blomberg, 1998; Muller, 2009; Schuler & Namioka, 1993). It is connected to ethnographic research because the "…value

[is] placed on participants' knowledge of their own practices" (Blomberg & Burrell, 2009, p. 88). PD has also proven to be of use "…as a way of jointly constructing with participants knowledge of local practices" (Blomberg & Burrell, 2009, p. 88). (Bødker, Kensing, & Simonsen, 2010; Crabtree, 1998)

When describing the relationship of PD to product design, Lanzara (1983), suggests that large-scale projects mostly involve resolving how conceptual and mental models translate into design work. In other words, designers are tasked with taking the input from many different user-based stakeholders and translating their collective feedback into definite and relevant problem statements to solve. As will be described later in this research effort, the ethnographic, PD-based methods used in this effort were essential to jointly construct many of the research artifacts with SMEs and to synthesize their various feedback into digestible problem statements.

2.11: Rapid Prototyping

Prototyping has existed for as long as mankind has been capable of using tools to create things: in physical and conceptual forms. With the advent of computer software, the creation of digital prototypes were added to our capabilities; things that don't exist in physical form, but are not simply abstract mental models either. They live somewhere between those two extremes. In terms of UCD, rapid prototyping methods include low-fidelity artifacts, like conceptual white board drawings of what a system could potentially look like, and more high-fidelity artifacts, like a basic web page that allows a user to complete a single task. These rapid prototyping methods are not meant to create a system that is 100% functional and ready for public release, but are designed to give the user a basic understanding of what something might look like and how it might operate in order to gather useful feedback throughout the technology development lifecycle.

In relation to this research effort, rapid prototyping methods are used to gather design feedback and requirements early in the technology development process. These mostly low- and medium-fidelity artifacts serve as disposable designs that take low levels of resources and require short schedules to finish. They can then be placed in front of SMEs and coupled with user-centered design methods in order to determine whether the prototype serves a purpose in aiding the expert to do their job.

2.12: Summary of Literature

Although MR technology has been a topic of study for decades, there is still much work to be done to better understand how to rapidly prototype mobile head-worn MR interface and interaction systems. Because MR headsets are becoming more consumer friendly, standards should be established in how to interact with these systems and how to best convey information to the user. One of the challenges of a novel technology, like that of MR, is that a multiplicity of related research domains and theories apply to the research effort. Basic human factors challenges exist in this domain, but the exploration of UCD methods will help to explain where specifically implemented processes are well-suited to the development of mobile head-worn MR interface and interaction systems in the first responder domain.

3: Objective 1 (O1): Information Requirements

The ultimate purpose of this study is to ***recommend rapid prototyping methods for mobile head-worn MR interfaces,*** using first responders (i.e., police, paramedics) as the user group of focus. The following Figure provides a visual reminder of the three research Objectives described in Chapter 1 and how this chapter fits into the overall effort:

Figure 6: Abbreviated Research Effort Overview

Along the course of identifying information requirements, this work has examined human factors research questions relevant to each objective. The following chapter describes O1 of the effort, which represents the first of the three main objectives. The stated research questions (RQs) of O1 are as follows:

- RQ1: What information do first responders expect to be available to them in a futuristic mobile HWD MR interface?
- RQ2: What information is most critical to the first responder to allow them to perform their job safely?

The first portion of this research effort began with a series of basic information-gathering methods that informed all of the design and prototype iterations going forward. By collecting both qualitative and quantitative user-centered data that supported each RQ, through the use of semi-structured interviews with subject matter experts (SMEs), researchers were able to better understand the domain of application for first responders (e.g., police and firemen) and create generic user-based stakeholder profiles, scenarios, and storyboards. The output of this applied method has provided the research team with an understanding of the landscape of how first responders perform their duties today at a micro-level and also how their specific tasks and responsibilities fit into the larger macro-level team in which they perform. This largely systems engineering understanding was vital to the creation of realistic user-based scenarios and storyboards to drive the effort forward. User requirements were also determined from the information gathered in semi-structured interviews. These user requirements were examined and summarized by researchers. They were then reviewed with SMEs to ensure an accurate representation of user groups and scenarios.

The following Figure provides an overview of O1, including the four Activities that were completed, the methods that were employed, and the research artifacts that were gathered throughout this portion of the research effort:

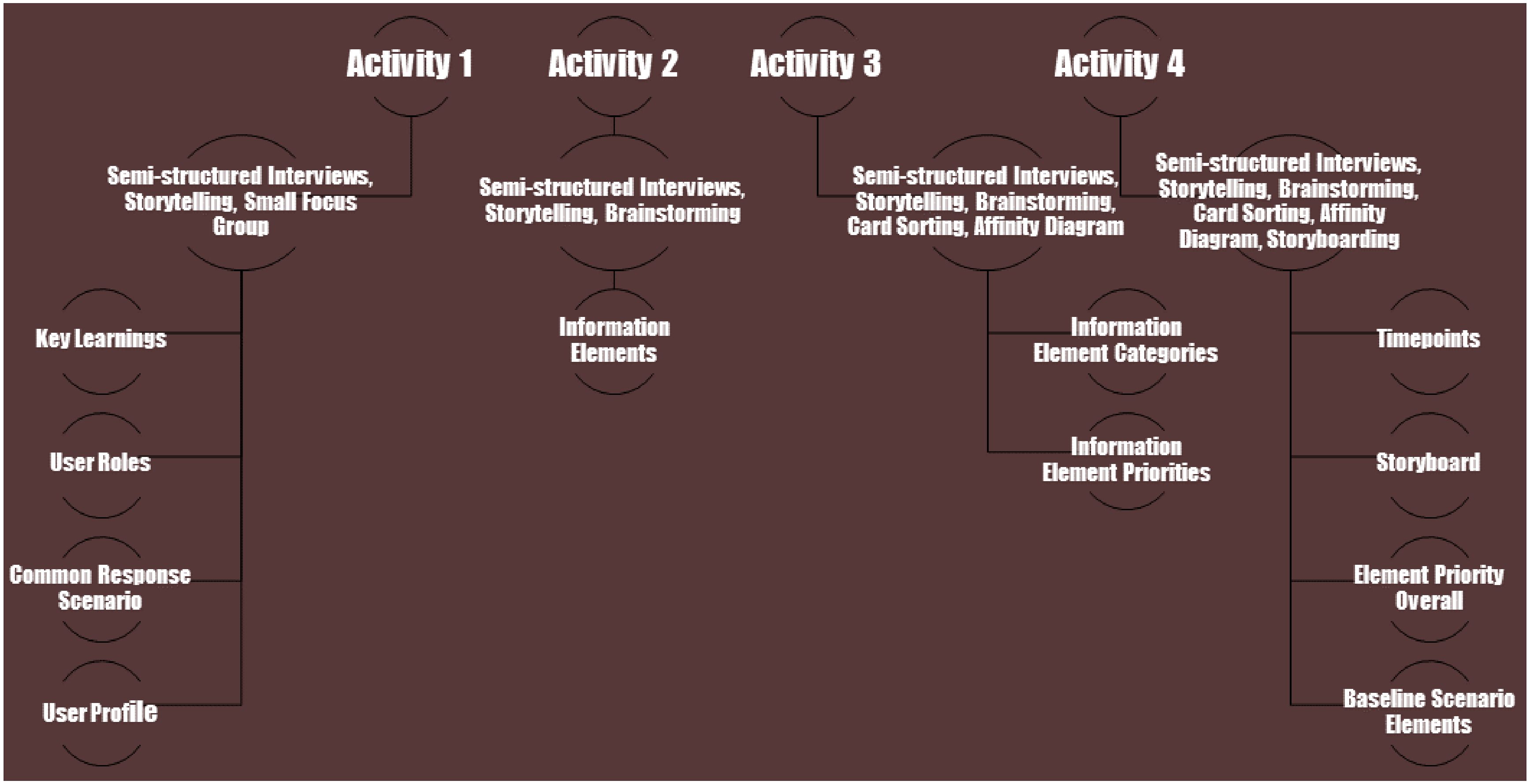


Figure 7: Objective 1 Overview

The rigor of the semi-structured interview and Participatory Design (PD)-based storytelling methods, coupled with the creative output of brainstorming, card sorting, affinity diagrams, and storyboarding methods, resulted in a high quality level of user-based feedback – in terms of detail and substance – in a logistically-convenient period of time (i.e., were short in their duration), in a rapidly iterative environment, because of immediate access to first responder participants.

3.1: Activity 1: Who, Why

The purpose of A1 was to better understand the domain of application for this research, including who the users were and why research in this domain would serve to further the academic progress of prototyping methods for 3D MR HWD systems. The following Figure provides a look at the user-centered design methods employed and the research artifacts constructed during A1:

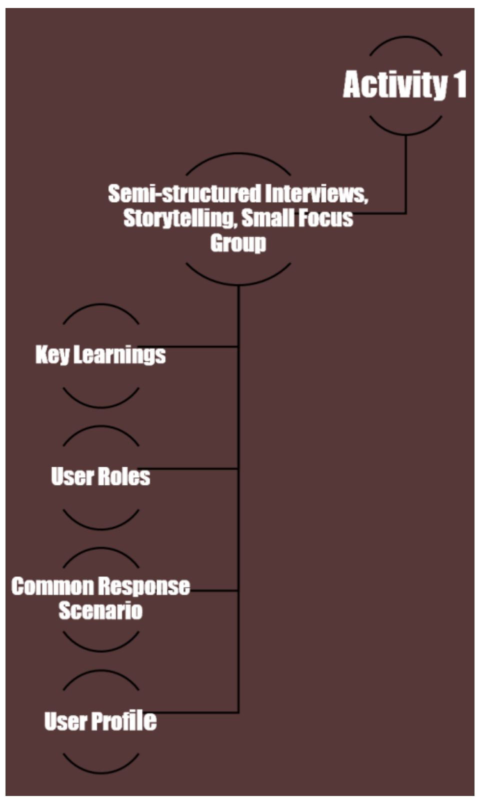


Figure 8: A1 Overview

3.1.1: Method

3.1.1.1: Participants

Initially, because this research effort involved a domain in which the research team had no expertise and very little professional interaction as well, a very wide net was cast in order to gather user-centered feedback from any and all available persons with any level of subject-matter expertise. This resulted in semi-structured interviews with a wide array of demographic participation; men and women, novice to expert, 18 to 60 years old, current and retired first responders from many different domains (e.g., military, medical, police, aviation). Participants were located in four different states across the United States – California, Connecticut, Florida, and Virginia – with experience at local, regional, and national agencies. Additionally, these participants performed their professional duties in myriad parts of the globe. 12 individual participants were interviewed and no monetary compensation was given for their time. The same group of 12 participants were recruited again to participate in a second iteration of the semi-structured interview process for this Activity.

3.1.1.2: Instruments

Handwritten notes were taken by the research team that were used later for post-interview analysis. No recordings were made during this Activity in order to allow participants to speak freely and confidentially.

3.1.1.3: Procedure

Each participant was asked a series of questions individually and in person in the form of a semi-structured interview. Each participant was seated across from the research team wherever they were located in their workplace (e.g., in their office, in their conference room) and were given a brief verbal description by the interviewer that the interview was completely confidential and simply served to give researchers an overview of the first responder domain as a whole. The following structured questions were asked of all participants:

- What is your last name?
- What type of job do you/did you do?
- What does a typical day look like for you?
- What do you like/dislike about your job?

During these interviews, the research team began to understand more about the first responder domain, including what types of tasks and processes are the most popular to perform, which are the most tedious, which are the easiest, etc. Based on what researchers perceived to be important to understanding the context of the domain, additional un-structured follow up questions were asked of the participants, such as:

- What technologies do you use in your work?
- How do you think futuristic technologies could help you be more safe at work and do your job better?

Additionally, when the interviewer perceived from verbal and non-verbal social cues that the participant had more to say about a given topic or subsequent clarification was necessary, further open-ended investigatory questions were asked, such as:

- Tell me more
- What else?
- Why is that?

A final question was asked when it appeared that the discussion had generally concluded:

- Is there anything else you have been thinking about during our discussion that I have not asked you about yet?

Once both parties were satisfied with the discussion, the interviewer asked for a referral on who they should speak with next, left the proximity of the current participant, and moved on to interview the next participant.

The same procedures were followed again for the second iteration of this Activity in terms of physical arrangement and location. Each semi-structured interview was individual in nature and emphasis was placed on having a brief interaction for this iteration. Researchers asked the following questions in this iteration:

- What is a common high-performing first responder's job title?
- How old are they?
- How long have they been doing their job?
- What types of emergency activities are most common to those specific jobs?
- Which of the common response scenarios are the most dangerous?

When necessary, additional follow-up questions were asked, but only to further explore more fine-grained details with the participant. To conclude this Activity, participants were shown a User Profile artifact in a small focus group and asked to come to a consensus on a final common response scenario that would apply to the ideal user profile.

3.1.1.4: Summary

No time limit was enforced in a participant's responses to any question and any question that the participant asked was answered by the interviewer immediately. No participant withdrew from any interview. Once the notes of all 12 participants were gathered together, a review of the written information and an analysis was performed in order to:

- Discover preliminary trends
- Determine which areas of the domain required further research and exploration
- Decide how many additional participant interviews would be required to understand the domain enough to progress in the subsequent design and prototyping efforts

Researchers attempted to limit the conversations of the second iteration of this Activity to 10 minutes or less. The final small group discussion of this Activity was an impromptu one held when three SMEs visited the research team and asked how things were progressing with the research effort. This small group meeting lasted no more than five minutes.

3.1.1.5: Results

12 first responder participants were interviewed individually in semi-structured interview sessions. Interview durations ranged from 30-60 minutes in range. The following Figure represents excerpts of two commonly iterated points from the notes taken during this exercise:

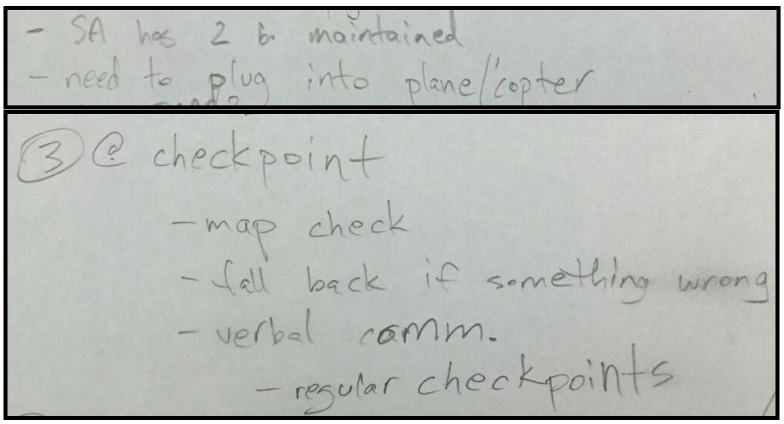


Figure 9: Note Excerpts from User Interviews

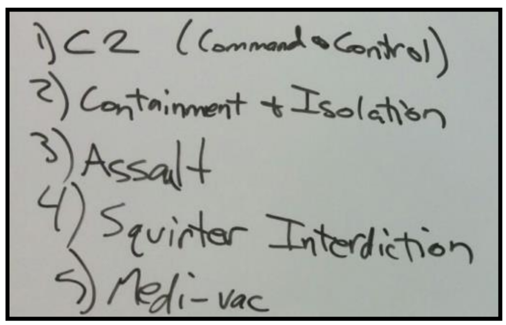
The following list summarizes the 12 key domain learnings from this Activity:

- Want higher-level understanding of area via maps
- Want to have conformal MR markers of all information all over the place
- Want to know where my friends are
- Want to know where the bad guys are
- Radio information
- Situational Awareness is paramount
- Don’t put too much information on my screen

- Want to clear/restore information on my screen easily
- Offload radio calls to text messages
- Customization is key
- Leave digital markers/pointers
- Video sharing is important

During the second iteration of this Activity, 12 additional semi-structured interview sessions were conducted. The names of various job titles and user roles were gathered and written on a white board, including the partial list shown in the following Figure:

Table 2: Team Roles White Board Session



The most common user roles from this list were used to create a cross-organization user role naming convention that applied to multiple first responder domains, as shown in the following Table:

Table 3: Most Common User Roles

Role #	Role Title	Role Responsibilities
1	Command & Control	Any leader on the ground.
2	Radio Operator	Asset coordination. Work with Role 1, talk with everyone
3	Perimeter Control	Containing the house and keeping everything else out
4	Assault	Clear and secure building
IN 50/50 SCENARIO:		
5	Pursuit	Chase the runners down
6	Medical emergency	Evacuate & move to safety
****Target user is for Role 4 primarily to increase safety level and situational awareness**		

The Assault role was identified as the primary focus for this research effort, defined as the first person to enter into a danger zone. Participants agreed that most high-performing first responders had the following characteristics as well:

- Approximately 10 years of experience
- Approximately 30 years old
- Lieutenant or Sergeant in rank (depending on specific organizational rank advancement structures)

Many response scenario suggestions were gathered from the first responders, but when the follow-up question was asked as to what was the most dangerous of the common responses, one scenario stood out immediately to all participants:

- Clearing a building of danger for my team (i.e., moving from room-to-room and ensuring no dangers are present)

These characteristics resulted in the following user profile, which addresses an ideal participant:

First Responder User Profile

Name: John Tentpeg, AKA "Duke"
Rank: Sergeant

SGT Tentpeg is a 30-year old male who enlisted in the military when he was 20 years old. With 10 years of experience, he has served his country around the world. As a soldier, he is most concerned with the dangers of urban warfare and is excited about the possibility of future technology systems that can offer him increased safety and situational awareness as he executes his future missions.

SGT Tentpeg is a 30-year old male who joined the police force when he was 20 years old. With 10 years of experience, he has served his community countless times. As a police officer, he is most concerned with the dangers of urban environments and is excited about the possibility of future technology systems that can offer him increased safety and situational awareness as he continues to serve and protect.

Figure 10: User Profile Example

The final data point for this Activity was developed in a small focus group meeting of three SMEs.

A review of the User Profile Example, along with the common response scenario, was performed. The first responder participants then verbally acknowledged that these research artifacts were accurate in their descriptions and useful to the research team and participants in framing future research progress.

3.1.1.6: Discussion

Preliminary analyses of this iteration of semi-structured interviews resulted in several important findings and insights. A wide array of information was gathered during these sessions. While this proved useful in understanding the domain of first responders in general, it did not provide consistent and specific feedback that would allow for the effort to progress with a targeted scope. In other words, there were no clear user trends or themes that would feed into the input of a rapid prototype design iteration after the first set of user interviews.

Specific first responder subject matter expertise was an essential influence to the content of the semi-structured interview feedback. Each first responder only felt qualified to comment on his or her specific job. Asking a patrol-level police officer about what a police chief might do in their work was perceived as outside the scope of their expertise. Asking a soldier about what a police officer might do in his work, although considered to be similar types of work as they are both part of the “security” subdomain of first responders, yielded little usable data in these interviews because they had no direct expertise in that job role. Even subdomains that share similar high-level tasks and knowledge were perceived to have a great deal of highly specialized information that was required to perform that job role. For instance, the civilian and military versions of the team roles described in “Table 3: Most Common User Roles” were only closely aligned when the specialization of that specific role was perceived to be at the same level of team organization. This equated to a SWAT-trained police officer and a Special Forces-level assault member being perceived to be generally equivalent in training and experience while a patrol-level police officer and basic infantry soldier were at a more generalized level of training and experience.

The actual scenario that was being responded to by the participant dramatically changed the content gathered during the interview. The processes and procedures that were followed during a response scenario were a direct result of the type of emergency that was in progress and varied widely depending on the emergency. For instance, a patrol exercise, where police or military personnel were traversing a region and looking for potential threats involved a completely separate series of actions and attitudes than when a civilian was identified as being in eminent danger and required immediate emergency support (e.g., domestic dispute, hostage scenario).

Because of the lack of commercially-available MR technology available today for first responders to use, it was essential to acknowledge initially in discussions and remind participants frequently that we were discussing scenarios that could potentially occur in the future (e.g., five or more years away). Without grounding discussions in the realm of a possible future, many first responders immediately dismissed the postulation that MR would be useful to them because current technological systems are perceived by first responder participants to be limited in their capabilities (e.g., field of view, portability, computing power).

The choice of a semi-structured interview was quickly identified as the information-gathering UCD method of choice for this Activity. Because the research team did not know what they would need to ask the participants, apart from a few basic boilerplate questions, it was essential to be prepared with something to ask the participant, but to allow the freedom in questioning and scenarios to investigate potentially important subjects during the interview process. Until a better understanding of both the content and context of the first responder domain was gained, a truly collaborative UCD-based communication structure was not possible between SMEs and the research team.

As a result of this widely varied user feedback of seemingly disparate information at times, researchers decided that an ideal user profile was needed to focus future work. This decision was further strengthened by the insight that many SMEs would ask the researchers for clarification on the exact circumstances of the response scenario (e.g., is a citizen injured? Is a vehicle involved? Is a team member in danger?). A second round of semi-structured interviews was necessary to refine the first responder scenario within A1 to specify a team member job title (e.g., police officer, paramedic) and common response scenarios (i.e., emergencies that were frequent and common) that could ground future research by gathering information that would be more unified in its content and guide future design and rapid prototyping decisions.

Due to several factors, including the lack of compensation given to participants and the interruptions that were occurring in the participants' workplace, the researchers decided that it was better to gather small sets of data more frequently with participants than to have longer duration interviews that were more in-depth. This was a result of attempting to schedule more formal times and durations for interviews with SMEs; which did not work well when attempted. Instead, when participants were asked, "can I ask you a couple questions real quick?," they were much more willing to stop their work and spend time providing short feedback. This was an important decision to this and future interactions with participants because participants could rely on quick iterations and low levels of time commitment required to provide useful feedback to researchers. In turn, the research team was able to gather the feedback required to make progress toward future information requirement, design, and prototype iterations later in the effort.

The introduction of the qualifying adjective "high-performing" to the first responder phraseology proved important to focus the mental models of participants. Researchers were not interested to know what untrained or novice first responders thought, but what the best and brightest represented as a whole. The experience, age, and rank advancement qualifications served to weed-out low-performing persons while the autonomy and leadership characteristics of the ideal first responder profile meant they required very little outside support for their job (e.g., they were leading and mentoring others). This narrowing of a user profile resulted in an ideal participant for the research team. As the profile was formed in real-time, participants were asked to verify that the researchers were interpreting their feedback correctly. This resulted in a profile that felt very "familiar" to the pool of participants. In statistical terms, this ideal user profile grouping was meant to represent one standard deviation in a bell curve of first responders.

The merging of the police officer and soldier job titles was made after verbal feedback from the participants in a small group setting; most of whom came from military and police backgrounds. The advanced experience and levels of expertise were what made a highly-capable team member from the civilian world generally equivalent to the military security domain; not just the job title they had.

The specific common response scenario (i.e., clearing a building of danger) was agreed upon to be a common task that presented the highest level of danger to security forces. This highly focused response scenario provided the context to future participants that would allow useful data to be gathered by the research team throughout the research effort.

The participants agreed that this same scenario was actually applicable to multiple other security-related domains (e.g., all law enforcement, security, protective, and military personnel around the world). In other words, A1 distilled not only an ideal user profile that was rooted in iterative UCD-based feedback, but also a specific common response scenario that would apply to more general domains of application, increasing the power of future design and rapid prototyping iterations by applying to a larger general population.

No major issues presented themselves during A1. This was largely due to the inexperience of the research team, however; researchers did not understand what they did not know until a post-analysis of each Activity was performed. Semi-structured interviews proved very effective in gathering initial data and understanding more about the first responder’s challenges in the workplace. Due to variability across the many tasks a first responder might participate in during a given work shift, the research team quickly discovered that a single user profile was essential to guiding the direction of the conversation. While this eliminates the consideration of outlier challenges to first responders, the need to focus on a single ideal user profile was determined to be more important at this stage of the research effort than enlarging the scope of the effort to include many different types of prototypical first responder profiles. After the completion of this activity, the research team did not regret this tradeoff in scope; it was necessary at the time to work within a set of given constraints (i.e., rapid iterations with short-duration scheduling and low investment of resources) in order to make progress toward future Activities and Objectives. Additionally, the common response scenario appropriately grounded the discussion of future MR experience design and rapid prototyping activities and iterations. Researchers now felt prepared to discuss what information was important to SMEs and how to best convey that information to them in a futuristic MR HWD experience.

3.2: Activity 2: What, How

While A1 helped to define who the targeted user was and why this research would contribute to the progress of rapid prototyping methods for 3D MR HWD systems, A2's purpose was to more fully understand what information was important to the user and how to best communicate that information to the user. The following Figure provides a look at the user-centered design methods employed and the research artifacts constructed during A2:

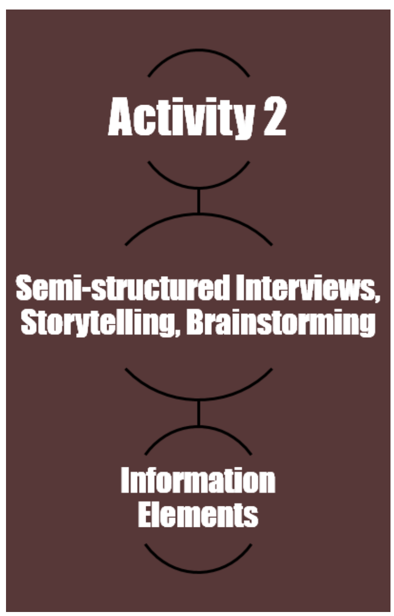


Figure 11 A2 Overview

With a general understanding of the research domain landscape and how the user currently performed their duties in the common response scenario, the research team made a UCD process plan for how to proceed with the effort. The following Figure shows the handwritten notes from this plan:

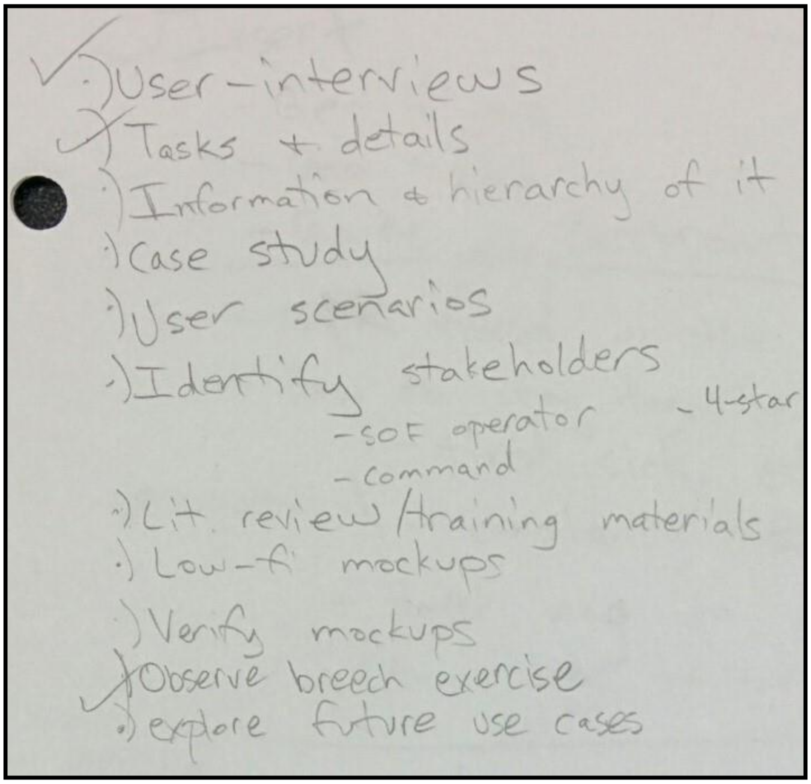


Figure 12: UCD Process Plan

This plan had been partially completed by the time it was digitized for the research effort and shows a general overview of what tasks were expected to be accomplished in the near future. With this plan in place, an enumeration of information elements was performed through the use of UCD-based semi-structured interviews and Participatory Design (PD)-based storytelling methods. The focus of these interviews was to explore and document the information that was currently on-hand to the SMEs in their workplace. The output of A2 provided the research team with a laundry list of the specific pieces of information gathered here. This includes information passed to the SME from other stakeholders (e.g., internal and external) via multiple modalities (e.g., communications on phone and radio, reference information from area maps, and specialty manuals). The storytelling method was especially useful in helping the SME to envision themselves in their duties while responding to an incident although they were in a mostly sterile laboratory environment (e.g., sitting in an office chair in a meeting room). Similarly, to the previous Activity's outcomes, these semi-structured interview and storytelling methods resulted in a high quality level of user-based feedback, in a logistically-convenient period of time, because of immediate access to SMEs.

With a greater understanding of the current state of information elements, the research team was then able to brainstorm with SMEs and compile a list of all the different pieces of information that could be perceived as useful to the SMEs in their current professional duties, regardless of what information was currently present during their professional day (i.e., the information currently on-hand to internal and external stakeholders). Semi-structured interviews and futuristic storytelling methods were used again to allow the research team to gather a "wish list" of information elements that could be available to users and external stakeholders in the future (i.e., the information that would be "nice to have") without the mental restrictions of currently-fielded technologies. This resulted in an expanded list of information elements that encompassed a more future-forward perspective – a place where current design constraints do not exist – so that future technology creation processes can be better prepared for changes in a more longitudinal system development lifecycle.

3.2.1: Method

3.2.1.1: Participants

Six total participants were recruited for this Activity. All six were a subset of the group of 12 that had been used in previous Activities. Three of them represented the ideal first responder user profile and regularly participated in the common response scenario, as detailed in A1. Two participants were first responders with five additional years of experience (i.e., 15 years instead of

10), but served in supportive roles to the common response scenario (i.e., they did not clear the buildings of danger themselves). These first responders had the same training to perform the common response scenario, but did not participate as frequently with that specific task as the ideal user profile group. The sixth participant was a second-level command structure participant who had additional years of experience and was serving a more administrative role to first responders when he was interviewed. This participant had performed the common response scenario many times in the earlier years of his career, but also had an understanding of higher-level processes and systems in the organizational hierarchy due to his longer service in the first responder community. The same six participants were then utilized in a second iteration of this Activity to gather more data.

3.2.1.2: Instruments

Once again, handwritten notes were taken by the research team that were used later for post-interview analysis. No recordings were made during A2 in order to allow participants to speak freely and confidentially. The enumerated list of information elements was later digitized into a spreadsheet format and standardized to combine the different phraseology used by each participant that was meant to represent the same individual information element. More details on these notes are provided in the Procedure section.

3.2.1.3: Procedure

The same procedures were followed for this Activity as the previous one in terms of physical arrangement and location. Each semi-structured interview was individual in nature and emphasis was placed on having a brief interaction for A2. Researchers began by grounding each participant in the ideal user profile and common response scenario (e.g., "You are a first responder who has been asked to clear a building of any dangers inside it"). PD-based storytelling methods assisted the research team in engaging the cognitive processes of the participant to mentally recall the common response scenario in their mind, although the interview was physically held within a seated office workspace. The participant was then asked to mentally picture themselves performing this task and to provide verbal feedback on what kinds of information they access and rely on to perform that specific task today:

- What kinds of information do you use today to do your job?
- What pieces of information do you utilize during the common response scenario?

The research team then feverishly wrote down every information element that was cited by the participant. The last three first responder participants (i.e., two support roles and one command role) were also asked to mentally explore the same scenario but from their native external support role. This list was also gathered by the research team.

The second purpose of this Activity was to understand how the information elements were currently conveyed to the participant in the common response scenario. The following questions were asked after the enumerated list was recorded:

- How are these information elements communicated to you today?

When necessary, additional follow-up questions were asked by the research team, especially when a new acronym or an unfamiliar term was used by the participants (e.g., “code 999,” “L.T.,” “Blue Force”).

In the second iteration of this Activity, the same initial procedures were followed as in the previous iteration. The participants were reminded of the ideal user profile and common response scenario, but this semi-structured interview stressed an additional component to the PD-based storytelling method: that the participant was now instructed to imagine themselves in the future (e.g., “You are a first responder living in the year 2020”). Within this future, the possibilities for information elements and interactions were endless as to the available technology to make it realistic. The participants were also told to provide verbal feedback on what kinds of information they would like to utilize in the future. The prompts given to the participants began as follows:

- If anything were possible, what kinds of information would you like to have access to?
- What pieces of information would you like to utilize during this scenario?

The research team recorded every information element that was cited by each participant. Like the previous Activity, the last three participants were also asked to enumerate lists of information elements from the perspectives of their supporting roles as well and those responses were recorded by hand.

The final purpose of this Activity was to also understand how the information elements could best be communicated to the future participant in this common response scenario:

- How would you like to be notified of an information element?
- What is the best way to communicate an information element to you?

When necessary, additional follow-up questions were asked by the researchers, but all interviews were kept brief to respect the participants' time.

3.2.1.4: Summary

No time limit was enforced during these interviews, but all interviews lasted less than 20 minutes each. Participants were given a little more time to respond to questions in the second iteration because this was an ideation exercise and required a greater level of mental creativity. The data was collected in the same manner and analyzed in the same fashion as in A1 by the research team.

3.2.1.5: Results

12 semi-structured interview sessions were conducted during A2 regarding the assault team member role. Additionally, six interviews were conducted regarding the support team member role. Approximately 30 different pieces of information were gathered that applied to the common response scenario, in addition to how they were communicated to the participant. The following is an example of information elements recorded during A2:

- Current time (via watch)
- Location of the emergency (via radio call)
- Who else is supporting me? (via mental recollection of who is working with me right now)

As mentioned previously, six of the semi-structured interviews were performed from supporting role perspectives of the common response scenario. These perspectives provided few additional information elements themselves, but were generally a more abstract view of a previously mentioned element (e.g., “what other teams are in place?,” instead of only “who else is on my team?”).

During the second iteration of A2, the same number and focus of semi-structured interviews were conducted in the futuristic information element domain. This provided an addition of 50 information elements to the original list that applied to the common response scenario, including the following example elements:

- MR markers
    - Where are my team members?
    - Where are the bad guys?
    - Where are we going?
    - Where have we been?

This resulted in an information list of 81 elements during A2. The following Table shows the entire list of information elements at this point in the research effort:

Table 4: Information Element List from A2

is radio operational	user temperature	rear view mirror
radio channel info	user blood glucose	health vitals of team members
comms rado, channel, callsign	remaining air supply	current altitude
fire coverage	whiteboarding sharing	enhanced vision perspective (e.g, thermal, night vision)
comms meshnet, BFT, threat location	video capture / playback / sharing	immediate threat locations
time of day	Video Down Link	shot sensor display
self injury indicator	video/images from other users	ground reference guide (as created during planning)
remaining power	video feeds from MY system	team tracking
ambient temperature	expected time until extraction	3D map of timelines and points
system temp	range and bearing to next objective	video feeds
diagnosis tests	expected time to target	breach points
access to reference documents	time since engaging target	incoming fire locations -- aggregate when appropriate
personnel files	self GPS location	supporting arms, assets
facilitiy files	user pulse	supporting arms, call signs
pictures of individuals of interest	compass heading	expected time to next objective (e.g., checkpoint, target)
operational docs	weather conditions	what parts of the system are malfunctioning
ammo inventory	mission map	weapon inventory
checklist	terrain data overlay	targeting graphics (for ranging and firing)
manifest	intelligence overlays	PPT brief created for planning, priortized slides
ambient air breathability	ocean conditions	distance to target
known danger areas	checkpoints	direction to target
alerts/notificaitons	nav routes	comms meshnet, team location range/bearing/elevation
user respiration	time to target	known geographic danger areas
self health vitals	image and video capture controls	sniper detection alert and location/bearing
remaining water/food supply	video and annotation UI	GPS location of team members
supporting arms, additional info	image capture UI	enviornmental and CBNRE detectors and notificaitons
self hydration indicator	speech-to-text autofill UI	speech-to text-transcriber for radio comms

All participants expressed a desire to have a larger perspective of the common response scenario communicated to them. Some of the suggestions included the following elements:

- A real-time bird's eye view of the entire area
  - Communicated via a 2D map
  - Communicated via a 3D map
  - Communicated via 3D MR-based markers

3.2.1.6: Discussion

The short-duration semi-structured interview method proved very useful in A2. Because of the good rapport that had been built between the research team and the participants during A1, each participant was willing to be interviewed briefly when asked. The decision to engage in more frequent, rapid iteration engagements resulted in no participant withdrawing from an interview nor declining a request to interview, although some interviews had to be postponed for a few minutes while the participant finished their current task.

The three support-role participants served to expand the scope of the information gathered and helped to bolster the feedback from the ideal first responder as it was related to the types of external information that were currently provided to them by supporting persons. The higher-level command participant also provided an important reminder that while the first responder in the common response scenario was only concerned with their own safety and the safety of whatever team members were performing the task with him, there could potentially be several teams that needed to be coordinated and organized. This systems engineering perspective was essential to consider when a larger building (e.g., a warehouse) or compound (e.g., a campus containing many buildings) was the location for the emergency response scenario.

Researchers also used this time to understand the way in which each piece of information was currently conveyed to the participant. These included multiple modalities (e.g., aural radio calls, visual laptop resources, cognitive recollection of past training, visual reference of manuals, haptic cues) and sources (e.g., within-team communication, external team support communication, command-structure guidance, interactions with citizens), which present additional human factors-related challenges (e.g., cognitive overload, ability to perceive speech, sensory overload).

This Activity did include a slightly larger aperture of research scope than the single ideal user profile in the common response scenario. Researchers feel this was necessary in order to ensure that all currently available information elements were enumerated for the ideal first responder, especially due to the smaller sample size of participants. Overall, these UCD semi-structured interview and PD-based storytelling methods resulted in a high quality level of user-based feedback, in a logistically-convenient period of time, because of immediate access to SME participants. While this slightly enlarged team role perspective could be considered to be scope creep, no research exploration is immune from such a pitfall and constantly vigilance is required to ensure resources are not unfairly focused in less fruitful areas as they relate to the overall research objectives.

No major issues presented themselves during A2 because of the presence of the ideal user profile, team roles, and common response scenario. And because this was simply an information gathering activity, there were no "wrong answers" to the questions asked. With the list of currently utilized information elements established, the research effort then reflected on the futuristic MR experience that would later be designed for the first responders. It was now time to place participants in an ideation exercise to make progress toward gathering what types of information could be useful in the future when a first responder would use an HWD in their daily work.

The use of handwritten notes proved very important to an iterative, rapid prototyping-style of information gathering. Because the research team did not rely on video or voice recording technologies, there was no minute-by-minute data to review after a participant interview was over; the research team had to be fully invested in the feedback that was given in the moment and quickly analyze that data after the interview was over. In some cases throughout O1, this meant immediately discussing trends with colleagues before making final conclusions. However, in later interactions with participants (e.g., O2 and O3), this meant that a more collaborative communication was occurring with the SMEs. Due to the gradually building rapport with participants, the SMEs perceived their feedback to be of great value to the overall research effort and were eager to assist the research team in designing and prototyping 3D interfaces for futuristic HWDs in a MR setting. It is important to note, however, that voice and video recording devices that can be analyzed in more detail and at a later date should be included in future design and prototype work.

Because this was a brainstorming exercise, it was important to allow participants time to mentally scour their cognitive processes for what they would perceive to be of use in the futuristic common response scenario. Depending on the duration of time that a participant had spent pondering the futuristic common response scenario, one could argue that a more formalized priming of participants (e.g., group-based brainstorming, presentations, instructions on the "art of the possible") could increase the quality and variety of feedback provided in A2, but such preparation requires a larger initial investment of time and resources from researchers in order to organize their participants. Due to the iterative, fast prototyping goals of this particular effort, such an investment in schedule and resources was not feasible at this stage of the research.

The possibility of a truly functional futuristic MR HWD technology presented the most exciting change to the participants during A2. While many information elements can be conveyed through multiple mediums (e.g., aural, haptic), first responders all remarked that they would prefer most of the list of elements to be conveyed visually as 3D assets presented on an HWD. According to the participants themselves, this was perceived to be the most effective way of communicating information and would relieve the cognitive strain that currently plagued them with other communication mediums; namely verbal messages via radio.

Due to the overwhelming number of information elements gathered throughout this Activity, there was a clear need for prioritization and categorization of the element list. Even with extensive training, the likelihood of any one first responder being able to utilize every information element that was requested would be highly improbable.

This ideation exercise was very helpful to brainstorm possible future functions of a 3D MR HWD system, but participants also used this session to give suggestions on future functions or capabilities of a possible system. While this was out of the scope of this specific information-element-based Activity, the suggestions were recorded and cataloged and, if anything, were an indication that the participants were solidly grounded in the futuristic PD-based storytelling scenario.

It was important to remind participants that current technologies and resources were irrelevant to the futuristic storytelling scenario; the researchers were concerned with future-proofing the list of information elements that could prove useful to any first responder. 3D MR HWDs are not currently fielded by first responders and will take many years to become ubiquitous in their jobs, therefore, because anything was conceptually possible during this ideation session, nothing was considered a "wrong answer."

Many of these types of elements, in terms of implementation, would require currently inaccessible external data sources and an untold number of sensors and communication networks to realize, which is an important design consideration for future research in this domain.

No major issues presented themselves during the execution of A2. This Activity did include a slightly larger aperture of research scope because it utilized the brainstorming/ideation exercise that was used to enumerate a more complete list of information elements (i.e., 81 elements at this stage of the research effort). This enlarging scope was determined to be essential to the research effort because it will likely be many years before a commercially available 3D MR HWD is fielded to first responders. If only currently available information elements were considered, a 3D MR interface would already be obsolete by the time it were actually built out by commercial vendors, due to the longitudinally-necessary scope, far-reaching schedule, and large investment of resources required to create such a system. Overall, these semi-structured interview and storytelling methods, combined with ideation methods, resulted in a high quality level of user-based feedback, in a logistically-convenient period of time, because of immediate access to SMEs and supporting personnel.

With an understanding of all the possible information elements that could prove useful to the first responder participant group, along with a perceived preference on how to communicate the elements via specific modalities, the research team was now prepared to move on to A3.

3.3: Activity 3: Priorities, Categories

While A2 helped to define what information elements were important to the user and how to best communicate them in a 3D MR HWD system, A3's purpose was to more fully understand where each information element fit into a categorized information hierarchy and which elements were most important to the user. The following Figure provides a look at the user-centered design methods employed and the research artifacts constructed during A3:

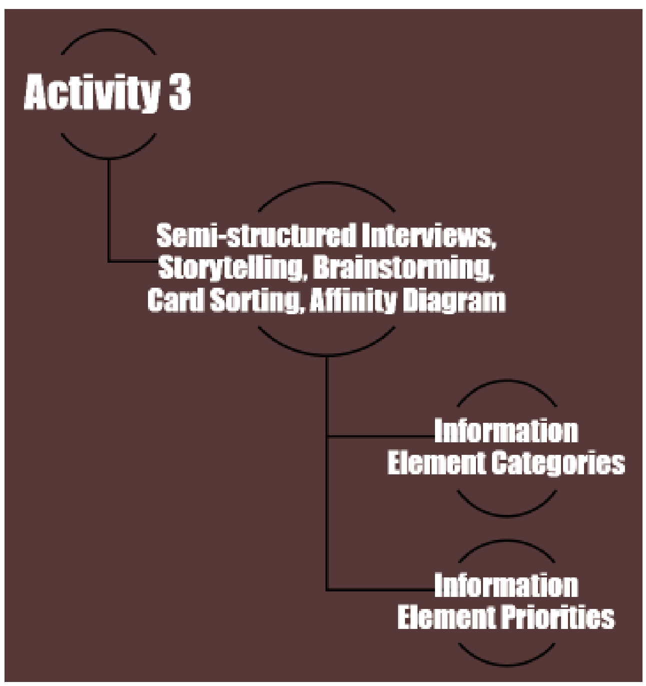


Figure 13: A3 Overview

User-centered semi-structured interviews and PD-based futuristic storytelling methods were coupled with a card sorting method; utilized by the research team in order to categorize and prioritize the list of previously defined information elements (n=81). A3 was designed to focus the research effort on how to organize the elements conceptually and what elements were perceived to be most useful to first responders during a common response scenario for mobile 3D MR HWD interfaces.

3.3.1: Method

3.3.1.1: Participants

Six first responders participated in A3. For more details on the participants, refer to previous Activities.

3.3.1.2: Instruments

As in previous Activities, handwritten notes were the recording method of choice for general participant feedback. However, photography was added as an instrument in A3 in order to simplify the recording of card sorting results. No video or audio recording devices were utilized. Each information element was written on a single 3x5" note card, which resulted in a collection of n=81 total individual card elements. A stack of sticky notes and a pen were also utilized in A3. After the card sorting exercises were completed by the participants, the research team digitized each participant's feedback into a spreadsheet format.

3.3.1.3: Procedure

The same locations were used in this Activity as the previous one. Participants were reminded of the ideal user profile (which was well-established by this time because each participant had been involved in multiple interview sessions previously) and the futuristic common response scenario of safely securing a building five years in the future. The PD-based storytelling method was used to mentally place the participant in this possible future where current technological limitations did not exist. Each participant was then handed the stack of information element cards and asked the following questions:

- While you are securing the building, what information elements are the most important for you to have in order to do your job?
- Of these elements, what is the priority of information to communicate to you?

The participant was then instructed to verbally explain their reasoning and prioritization; these notes were recorded by the research team and a photograph was taken of each sorting result. Once this sorting exercise was completed, the information element card deck was gathered together and a second version of the sorting exercise was performed. The participant was then instructed, using the same scenario and user profile, to take the card deck again and do the following:

- Organize the elements into meaningful groupings and categories
- Give each category a name that describes what that grouping represents to you

The participants then grouped the elements into categories and used the sticky notes and pen to name each category. The results were photographed and additional notes were taken by the researchers. When necessary, additional follow-up questions were asked by the researchers. Each participant interaction with the research team was kept to a minimum.

The research team performed an analysis of the collected card sort data and categorized the data according to their understanding of the domain and relevant notes/context gathered during this effort. This was performed because of the disparate categorical information gathered during the participant interviews. A final six-category organization was presented during a small collaborative design focus group session with three first responders. This resulted in a 13-category listing of information elements that was approved by the focus group participants as an appropriate information organization structure for A3.

3.3.1.4: Summary

No time limit was enforced for this exercise, although the range of time used in this experiment was between 10 and 30 minutes. The data was digitized and analyzed by the research team immediately following each participant's responses. A six-category information organization was developed by the research team, but this was quickly iterated into a 13-category organizational structure after feedback from participants.

3.3.1.5: Results

A total of six semi-structured interviews utilizing a card sorting method were conducted during A3. This produced twelve total card sorting results, which can be described as:

- Six ***priority***-organized information element sortings
- Six ***category***-organized information element sortings

An example of an individual card sort result is shown in the following Figure:

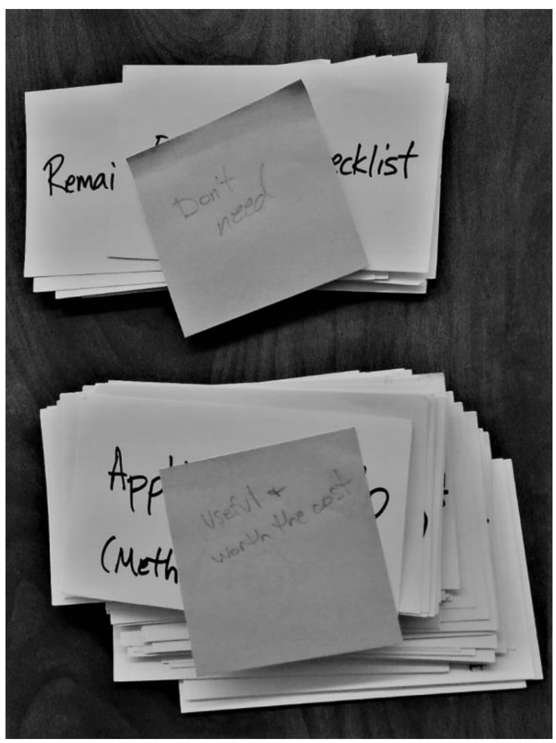


Figure 14: Card Sorting Example 01

This participant chose to categorize his elements into two simple stacks:

- Don’t need
- Useful and worth the cost

In contrast, another participant required a very limited set of information to be displayed on his futuristic 3D MR HWD, as shown in the following Figure:

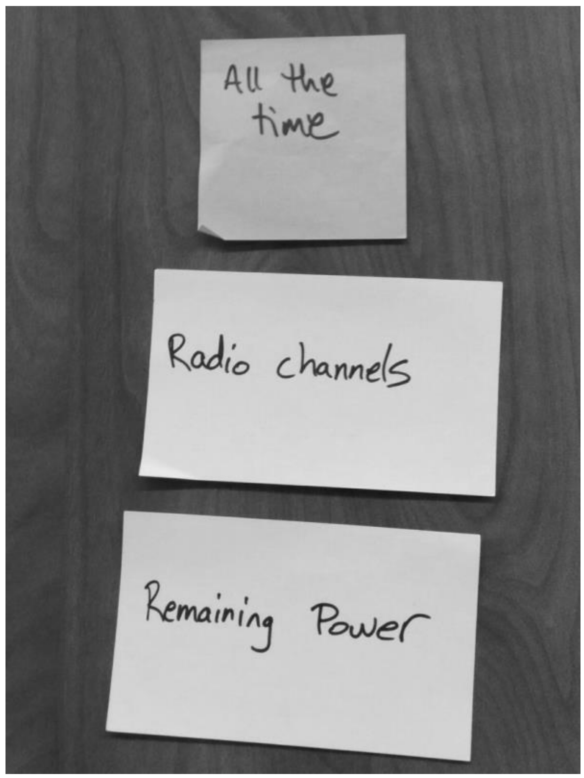


Figure 15: Card Sorting Example 02

This second example shows that this participant only wanted two information elements shown to him all the time: “radio channels” and “remaining power.” The rest of the element list would be turned on and off as needed. The following Table displays the number of categories that each participant chose during this exercise:

Table 5: Number of User Categories

Participant #	# Categories
1	8
2	6
3	10
4	11
5	13
6	2

Due to the large variation in participant-defined categories, the last three participants were informally questioned as to why their results were so different than that of their peers. Their responses were recorded and analyzed. The research team then attempted to categorize the elements themselves and developed the following information structure based on previous user feedback, as shown in the following Table:

Table 6: Research Team Categorization of Information Elements

Enhanced Vision	Navigation	Weapons	Mission Checklists	Notifications	Capabilities
Blue force tracking	Bearing	In-team support	Things to do	Ambient air breathability	Application I/O (Method TBD)
Call sign	Checkpoints	Fire coverage	Fire coverage	CBRNE Alerts/Notifications	Breach points highlighted
Team location	Direction to checkpoint	Range	Maps	Danger sensor display	Controlling system (Method TBD)
External supports	Direction to target	Targeting system	Preparation info	Team vitals	Data capture (of anything)
Enhanced Vision	Distance to checkpoint	Weapon System	PPTs	In-team support	Data review (of anything)
Surveillance Capability	Distance to target		Processing kit	Incoming fire locations	Data sharing (of anything)
Maps	Elevation			Known danger areas	Things to do
Range	Maps			Ocean conditions	External supports
Rear view mirror	Heading			Operator pulse	map elements highlighted
Incoming fire locations	Lat/Long			Outside Temp	Preparation for mission
Red force detection	Nav. Routes			Radio Channels	Meshnet
Threat location	Range			Radio Working?	Rear view mirror
	Terrain Maps			Remaining air supply	Speech-to-text & text-to-speech
				Remaining food supply	
				Remaining power	
				Remaining water	
				Self injury indicator - am I shot?	
				system OK?	
				diagnosis tests	
				system temp	
				Threat location	
				Time of day	
				Time since engaging target	
				Time to checkpoint	
				Time to target	
				Time until extraction	
				Vital signs	
				Weather conditions	

Key of Indicators:
Danger
Navigation
Friends
Weapons
Data Files
Etc.

This six-category organization was shown to a small focus group of three first responders. They critiqued the researcher-organized groupings and arrived at the following 13-category organization of how the information elements could best be structured for future design Activities:

Table 7: 13-Category Information List

Blue Status
Call sign
Team location
External support
Blue force tracking
Fire coverage
In-team support

Threat Status
Threat location
Incoming fire locations
New dangers identified
Known danger areas
Surveillance
Red force detection

Green Status
Friends
Surveillance

White Status
School/Church/etc. Identification
time of day milestones
Friendly Forces
Surveillance

Enhanced Vision
Enhanced Vision
Maps
Range
Rear view mirror
Long-range vision
Breach points highlighted
Surveillance

System Status
Radio Working?
Remaining air supply
Remaining food supply
Remaining power
Remaining water
am I hurt?
system OK?
diagnosis tests
system temp

Weather
air quality
Ocean conditions
Outside Temp
Weather

User Status
Health of team
Operator pulse
air supply
food supply
water
am I shot?
Vital signs

Data review
Planning materials
Preparation for mission
checklist
Nav. Routes

Weapons
Range
Targeting system
Weapon System

Navigation
Bearing
Checkpoints
Direction to checkpoint
Direction to target
Distance to checkpoint
Distance to target
Elevation
Maps
Heading
Nav. Routes
Range

Comms
Radio Channels
Radio Working?
Chat Window
Speech-to-text

Mission Status
Checklist
Nav. Routes
Surveillance
Time to checkpoint
Time to target
Time until extraction
Time of day
Breach points highlighted

3.3.1.6: Discussion

It was important for participant interaction to be kept at a minimum. The prompts for this exercise were very ambiguous and each participant was actually rather uncomfortable with the lack of structure and guidance in the prompts that were given to them. However, in order to not influence the later participants' responses, the same testing protocol was followed for each first responder card sorting session.

As the results were analyzed by the research team, there was no obvious consistency in the responses. Because the responses were scrutinized after each card sorting session, it became more apparent as the information-gathering process progressed that something was not quite right. This led to three informal investigative conversations with the latter three participants after their sorting exercise was finished and their responses were recorded in order to understand why there was so much variability in participant responses as to both the priority of elements (e.g., which were most important to display to the user) and the categories of organization (e.g., conceptual information organizations). These participants were shown the results of their peers and questioned as to why each first responder had dramatically unique responses. These final respondents indicated that the exact context of the common response scenario was still too vague for a proper card sorting exercise to yield the consistent results sought by the research team.

Although the resultant priority and category information gathered in A3 could largely be considered not useful to the overall RQs of O1, the realization of this additional required context marked a pivotal point in the research effort. As will be described later in this paper, without this important realization at the end of A3, it is expected that future Activities would have continued to experience similar data variation and more time would have been lost in progressing the design and rapid prototyping of the overall effort.

3.4: Activity 4: Storyboard Context

While A3 helped to categorize, organize, and prioritize the information element list according to the perceived needs of the user, A4's purpose was to more fully understand the context of the common response scenario and resolve the data variation issues encountered during A3. The following Figure provides a look at the user-centered design methods employed and the research artifacts constructed during A4:

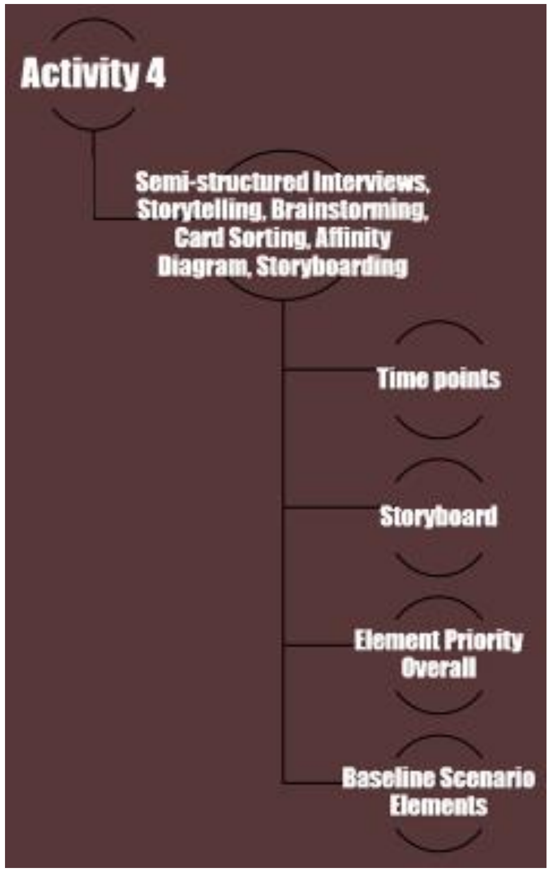


Figure 16: A4 Overview

With this previously missing context in mind, a four-participant user group was gathered together in A4 to discover the additional detail that was absent from the card sorting exercise of A3. A4 would explore the common response scenario in more detail and then execute a second iteration of the cord sorting exercise that was performed in A3 within this newly detailed common response scenario.

3.4.1: Method

3.4.1.1: Participants

Four participants from the ideal user profile group were gathered together for a small focus group collaborative design session. Seven first responder participants were utilized in A4 for a second iteration of the card sorting exercise used in A3. Six participants were also used for a small focus group to verify the findings of A4. For more details on the participants, see the descriptions in the previous Activities.

3.4.1.2: Instruments

A white board and marker were used for the small focus group exercise. Notes were made on the white board in collaboration with the first responder participants. The white board notes were captured via photograph and stored for future use in this research effort. An artist was also utilized during A4 to better convey the concepts being described by the research team and first responder participants. The set of 3x5" index cards from A3 that enumerated the list of information elements was also used in the A4 card sorting session. Handwritten notes, as needed, provided additional post-experiment analysis material. After each card sorting exercise, the results were photographed for later review by the research team. Each result was also digitized into a spreadsheet format.

3.4.1.3: Procedure

The participants were gathered around a white board in a common workspace. Researchers explained that the responses gathered in the last card sorting exercise did not provide a consistent picture of the priority and categorization that should be considered in a futuristic 3D MR HWD system design. Researchers perceived that this was due to a lack of the exact context and a level of specificity that was necessary to add to the common response scenario. As a result, these four participants were gathered together to help add additional context to the clearing a building of danger storyboard. The research team then asked:

- What is missing from the scenario description?

- What context is not being discussed in the questions we are asking you?
- Why does everyone have such different responses?

The research team then listened and wrote the notes of this discussion on the board. Five time points were standardized and turned into a storyboard format. Seven first responders were interviewed for a final time using the additional context provided by the five time points of the common response scenario. The testing protocol script for A4 is shown in the following Figure:

Confidential Demographics

- Date & Time?
- First Name?
- Where do you serve?
- Years Experience?

Timepoint 1:

You are a first responder on your way to a secure a building. You are the first person responding at the emergency. You are moving toward the target location.

- What categories of information are most important for you to know?
- What are the top 3 pieces of information that are most important to you at that moment?

Timepoint 2:

You arrive at your target and are getting ready to enter the building.

- What categories of information are most important for you to know?
- What are the top 3 pieces of information that are most important to you at that moment?

Timepoint 3:

You entered the building and are clearing rooms inside it.

- What categories of information are most important for you to know?
- What are the top 3 pieces of information that are most important to you at that moment?

Timepoint 4:

You cleared the building and everyone is safe.

- What categories of information are most important for you to know?
- What are the top 3 pieces of information that are most important to you at that moment?

Timepoint 5:

You finish at the building and it's time to go home.

- What categories of information are most important for you to know?
- What are the top 3 pieces of information that are most important to you at that moment?

FOLLOW-UP:

- Are there any pieces of information that you want to see all of the time?
- Is there anything else you would like to share with us?

Figure 17: Testing Protocol Script

Notes were taken during A4 and card sorting responses recorded via photograph. The research team analyzed each respondent's data immediately following their interview to determine whether the data inconsistency issues from the previous card sorting session had been resolved (e.g., disparate category/no common priority of element organization across participants). Each response was digitized into a spreadsheet format and scrutinized by the research team in terms of both an individual participant's responses (e.g., the most common elements across time points) and an individual sub-section of time's responses (e.g., the most common elements across participants in a single time point). It was assumed throughout this card sorting exercise that the response scenario went smoothly; no emergencies were experienced and no contingencies were needed that would alter the traditional first responder reaction.

A small focus group was held with six first responders, during which all six participants came to an agreement that the stated information elements, arranged according to the time points of the common response scenario (i.e., securing a house), were a good standard with which a future first responder participant could be introduced to a 3D MR HWD system.

3.4.1.4: Summary

No time limit was enforced for this Activity, although the range of time used in the experiment was between 10 and 30 minutes. The data was digitized and analyzed by the research team immediately following the small focus group meetings and each card sorting exercise.

3.4.1.5: Results

Participants provided the details of the different chronological segments involved in the common response scenario. Different names for the segments were written on a white board and then standardized into the following five time points:

1. Traveling to Building
2. Preparing to Enter
3. Securing Building
4. Processing Building
5. Going Home

These time points were formalized into a storyboard format that is shown in the following Figure:

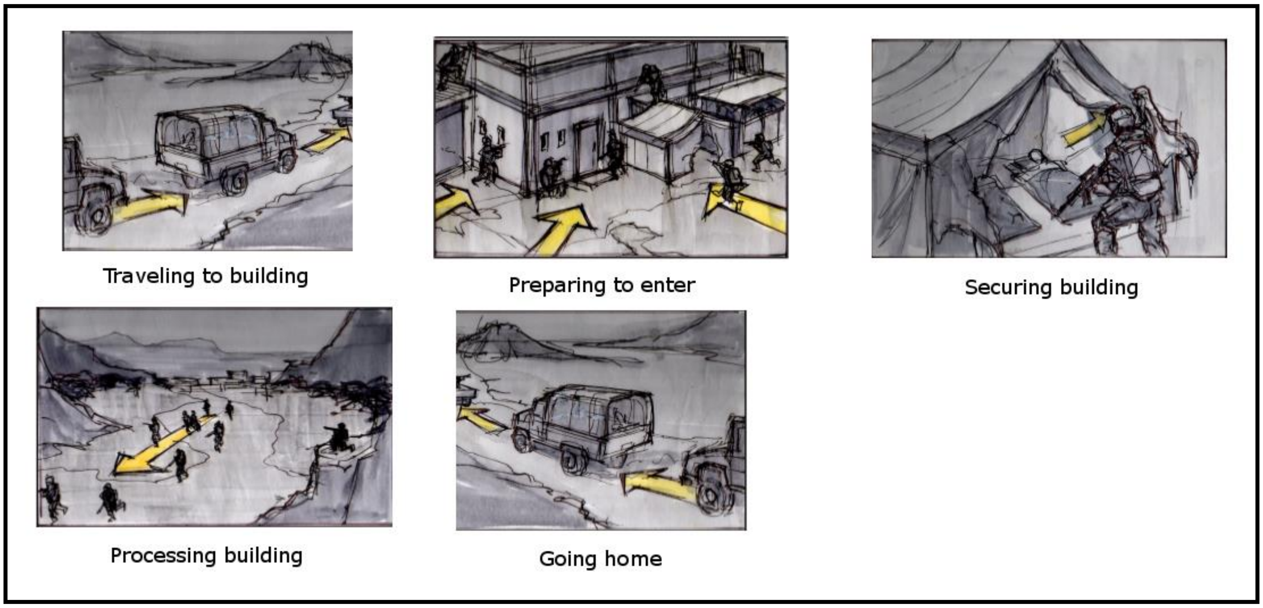


Figure 18: Storyboard with Five Time Points

The research team collaborated with an artist to develop the storyboard research artifact due to a lack of such creative talent within the research team itself. But the direction for what was to be drawn and the message that was to be communicated was under the immediate leadership of the researchers and first responders at all times. This storyboard was produced in a small focus group setting.

For the card sorting exercise, each participant was seated at their workspace. The same PD-based storytelling method was utilized in A4 as in previous Activities to properly place the mind of the participant in the appropriate context to gather relevant feedback. Now the first responder was asked to stop at each time point, mentally reflect, and organize the cards according to the testing protocol script, as was shown in the Procedure section of A4.

One example of the notes taken during one participant's interview is shown in the following

Figure:

5/22 @ 3:14 pm

[redacted]

Timepoint 1: moving towards location

- Navigation, Threat Status, Mission Status
  - Navigation and threat status shift in priority throughout approach
- Obstacles (tied to route), threat changes from Pre-Brief, FF status (everyone on the ground?)

Timepoint 2: getting ready to enter

- Threat Status, Blue Status, Mission Status
- changes to threat, ready for entrance (everyone in position?)

Timepoint 3: clearing rooms in the house

- Threat Status, Blue Status, Mission Status
- Room/floor/sector/building secure, casualties

Timepoint 4: you have successfully cleared house

- data capture, blue status, process scene
- who is here?, what's everyone doing? Everyone save?

Timepoint 5: finish and go home

- Threat Status, Navigation, Data sharing
  - As before, Navigation gains priority with distance from scene
  - Data share is important before next high risk situation
- changes to route, changes to threat, team status

Figure 19: Response Example from Second Card Sort

The demographic information for this participant is redacted from this research artifact in order to provide the necessary identity confidentiality. It was assumed throughout this card sorting exercise that the response scenario went smoothly; no emergencies were had and no contingencies were exercised. Each interview resulted in a list of the information element priorities and categories for each participant. A summary table was created that shows a final 17-category information organization of data by the first responder test group (i.e., the row headings) compared to the five scenario time points (i.e., the column headings) of this final card sorting exercise below:

Table 8: Summary of Categories Utilized in Common Response Scenario

Categories of Information	Scenario Time Points				
	1: moving towards target	2: preparing to enter	3: clearing rooms	4: rooms-cleared	5: finish and go home
Blue Status	3	5	5	4	4
Threat Status	4	5	5	0	3
Weapons	1	1	3	0	1
Navigation	7	0	0	1	5
Mission Status	2	4	1	3	3
Comms	2	2	2	1	3
Data Sharing	0	0	0	2	1
Processing	0	0	0	6	0
Questioning	0	0	0	5	0
System Status	1	1	3	0	0
White Status	0	1	0	0	0
Green Status	1	1	0	1	2
Data Capture	0	0	0	1	0
Data Review	1	0	0	1	0
User Status	0	0	1	0	0
Weather	0	0	0	0	0
Enhanced Vision	2	1	1	0	0

Key:
Most Common
Moderately Common
Least Common

This Table shows that when this analysis was finished, very clear information element priority trends emerged across the entire user group. The categories and time point pairings highlighted in green represented the most common categories identified as important to display at that point in the response scenario, while yellow highlights represented moderately common categories and red highlights indicated the least common categories needed at that time. It is also important to note that four additional categories of information were described by participants after the end of A4, expanding the previously used 13-category organization from A3. The additional four categories represented generally repetitive sub-sets of information and were typically the result of a few participants' personal preferences in information architecture and hierarchical organization (i.e., "Questioning" was one user's sub-category for "Processing")

The following Table shows a visual representation of the average number of priority elements gathered from this experiment, broken down according to response scenario time points:

Table 9: Average Priority Element Breakdown

	Response Scenario Sub-Sections					
	Traveling to House	**Preparing to Enter House**	**Securing House**	**Processing House**	**Returning to station**	
Average Number of Priority Elements	5	5	2	5	5	
Participants	7	7	7	7	7	
Average Total of Priority Elements	30	30	12	30	30	
					Total Average Priority Elements	**132**

132 useful information element data points were gathered during this second card sorting exercise, with an average of five elements displayed during time points one, two, four, and five. Time point three had an average of two elements displayed, also across all seven participants.

A final list of the most important individual data elements (not just categories) to display at each time point of the common response scenario for the assault team member role was created and verified with first responders in a small focus group to develop a final baseline information recommendation. All seven participants agreed that these information elements, arranged according to the time points of the common response scenario, were a good standard with which a future user could be introduced to a 3D MR HWD system. This basic information organization was termed the “training HUD.” Throughout this discussion, it became clear that any user would need to be allowed to fully customize their MR interface preferences; in both the number of information elements that would be communicated and the method of communication (e.g., visual, aural, haptic). The “training HUD” list of critical elements, organized according to scenario time points is shown in the following Table:

Table 10: Common Critical Elements, Arranged by Time Point, AKA, the “Training HUD”

Sub-section:	**1**	**2**	**3**	**4**	**5**
Elements Included:	Compass	Compass	Compass	Compass	Compass
	Time	Time	Gaze	Time	Time
	Radios	Radios		Radios	Radios
	Messages	Messages		Messages	Messages
	Gaze	Gaze		Gaze	Gaze

This more restrictive storytelling exercise typically resulted in a card sort of less than eight elements for each time point, prioritized by criticality, with the remainder of the information element list discarded at that moment in experiential storytelling time.

3.4.1.6: Discussion

Although the responses gathered during A3 indicated that the common response scenario did not include enough detail in order to gather the level of comprehensive card sorting feedback that the research team desired, it was important to state this sentiment to the first responders in the small focus group information gathering session to verify that the research team’s perception of their data-based shortcoming was indicative of reality. Once confirmed, this brought the research team and participants into a common baseline understanding that proved essential to future research progress.

This discussion revealed that the first responders had all been mentally organizing the common response scenario subconsciously into five distinctive time points. While this was intuitive knowledge to the participants, the research team did not clearly understand the existence of this chronology. It is unknown why this level of detail was never clearly identified by the research team earlier in the process. Perhaps ignorance within the research team as to the real-world expertise of the first responder domain was the culprit of this oversight. Or perhaps the priority placed on rapid iteration and development of the research effort resulted in a lack of the in-depth data necessary to properly understand the first responder domain for this research application.

A formal storyboard was useful to bounding the mental progress of first responders in the common response scenario. Apart from verbal communication and descriptions from the research team, this visual representation created an even more realistic mental model for first responders to describe their needs.

The context of these time points was a critical turning point in this research effort. While the results of the first iteration of the card sorting exercise were erratic and inconclusive, the injection of the five-time-point-chronology suddenly brought a clarity of purpose to the participants and resulted in a baseline of the categories in the second iteration of the card sorting exercise where participants came to an almost immediate consensus on the most important information to be used in the common response scenario.

Participants also began to use similar category naming conventions, even without the intervention of their peers during the card sorting exercise (i.e., each interview was performed individually and without external collaboration). This was partially a result of the adoptive categories of previous research iterations, but was also indicative of a “melding of the minds” in the progress of the first responder participant pool overall and the understanding of the domain of application within the research team. Due to the common workspace used during O1, it is possible that participants were having informal discussions about our experiments with each other on their own time, but it is highly unlikely that these “water cooler” conversations were having a large impact on this consensus-building user feedback data set. This was largely due to the fact that the participants never had any formal meetings or discussions on the experimental treatments of this effort outside of the small group collaboration sessions held with the research team.

The prioritized information elements that participants suggested should be included on a 3D MR HWD system of the future also rose to a level of consensus at the same time. It is important to note that it was assumed throughout this card sorting exercise that the response scenario went smoothly; no emergencies were had and no contingencies were needed. This allowed the participant to focus on what information elements were truly critical to accomplishing their task at that moment in time, without having to mentally process all the “what ifs” that accompany this typically dangerous and highly unpredictable operational scenario.

The common response scenario time points were integral to discussing all future experiments throughout this research effort, especially as related to later research phases. It is thought likely that without this time point differentiation realization early on in the research process, future design and prototype sessions would have largely failed because the context of the specific moments of time were critical second-nature types of cognitive context that were deeply engrained in each participant both mentally and perceptually. And without this seemingly small detail, participant responses would likely have continued to be conflicting among the users and unpredictable for the research team to gather.

Although each information element was provided by a first responder as a desired future capability of the 3D MR HWD system, there were simply too many cards to sort through. The cognitive overload from a stack of 81 information elements meant that no single participant could recall what the entire list consisted of in the first place; there were too many variables to consider depending on how the response scenario played out.

The final brief small focus group interaction with the first responder participants in A4 was important to double-check the work of the research team. Finding the six participants of this final review to be in agreement with the organization of the most critical information elements at each of the five time points of the common response scenario meant that a clear consensus had been found across the first responder test population.

Overall, these semi-structured interview and PD-based storytelling methods, coupled with brainstorming, card sorting, and storyboarding methods, resulted in a high quality level of user-based feedback – in terms of detail and substance – in a logistically-convenient period of time (i.e., were short in their duration), in a rapidly iterative environment, because of immediate access to first responder participants.

4: Objective 2 (O2): User Interfaces

The ultimate purpose of this study is to ***recommend rapid prototyping methods for mobile head-worn MR interfaces,*** using first responders (e.g., police, paramedics) as the user group of focus. For the benefit of the reader, the following Figure communicates the three overall Objectives of this effort. This chapter will discuss O2. See previous chapters for more details related to this Figure:

Figure 20: Abbreviated Research Effort Overview

Along the course of interface development, this work will continue to examine human factors research questions relevant to each objective. The following chapter describes O2 of the effort, which encompasses a description of rapid design and prototype user interface development experimentation. O2 answers RQ3 and RQ4, while O3 includes the final research work that answers RQ5 and RQ6. The stated RQs of O2 are as follows:

- RQ3: How can critical interface information be communicated to first responders utilizing multiple modalities (e.g., visual, aural, haptic) of notification?
- RQ4: How does the context in which a user experiences a prototype effect the interface feedback they provide?

Focusing on a methods-first approach, appropriate detail is given related to the execution of each Activity within this chapter so that relevant lessons learned can be garnered from each research experiment. By collecting both qualitative and quantitative user-centered design and prototype data that support each of the RQs related to O2, researchers were able to better understand how user interfaces could best be implemented within the first responder domain. Through the use of semi-structured interviews with SMEs, PD-based storytelling, small collaborative focus groups, and brainstorming ideation methods, the necessary qualities of an overall positive user interface within this highly specialized domain were better understood and can be used to inform future research applications in this area of interest.

The following Figure depicts an overview of the each Activity described in O2, along with the methods that were used during and the research artifact outputs:

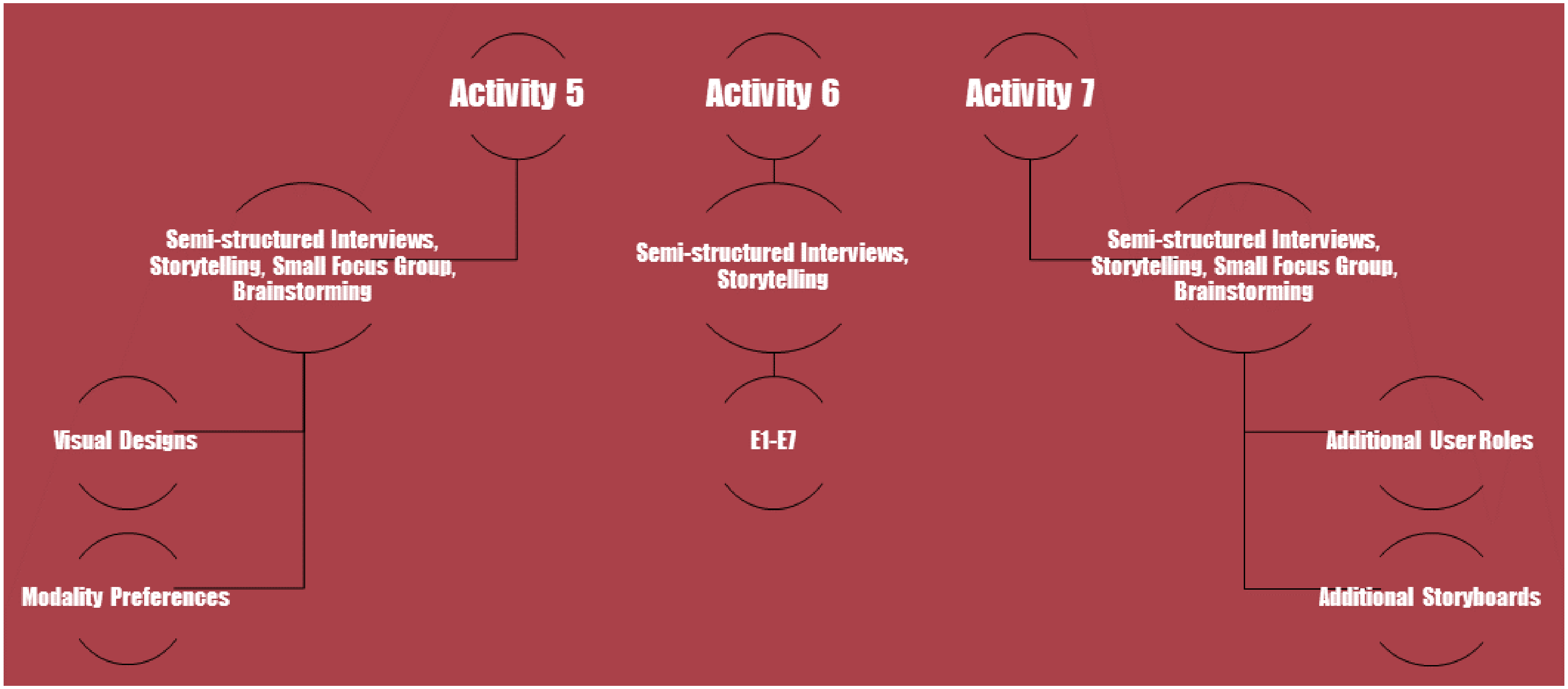


Figure 21: Objective 2 Overview

Similar to the discussion in the previous chapter, this chapter includes three unique Activities (i.e., A5, A6, A7) that comprise the work of O2. With the completion of the initial information requirements gathering in O1, the research team performed a series of iterative rapid design and prototype phases during which 3D MR HWD system user interfaces were explored. O2 was informed by the results of O1. A post hoc analysis of the methods and tools utilized during O2 will be described later for this portion of the research effort.

4.1: Activity 5: Interface Design

While A1-4 within O1 helped to determine the information requirements of the users and how best to communicate the highest priority elements in the common response scenario, no designs had yet been formalized. A5's purpose was to create user-centered design documentation and concepts that could be implemented into a user interface prototype and tested with first responders. The following Figure provides a look at the user-centered design methods employed in A5, along with the research artifact outputs:

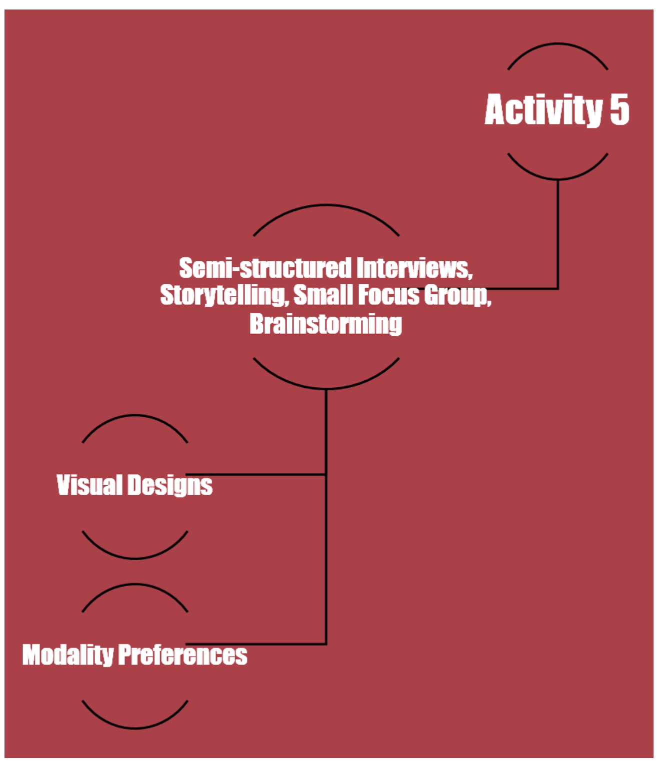


Figure 22: A5 Overview

4.1.1: Method

4.1.1.1: Participants

12 first responder participants were interviewed individually in a semi-structured interview session during A5. Six of them fit within the ideal user profile described in O1. The participants represented local and national agencies. No monetary compensation was given for any of the interactions described in this paper. Six of the participants (i.e., three of the ideal user group and three additional first responders) had previous experience with the research team in previous Activities, while the remaining participants were new recruits. See O1 for additional details on these participant groups. A small focus group was held with three first responders.

4.1.1.2: Instruments

White board sketches, digital sketching devices (e.g., touch-based tablets), paper drawings, handwritten notes, photography, and sticky notes were all utilized during A5 to record participant feedback. These digital and physical artifacts were recorded and archived and were used later for post hoc analysis and discussion. The storyboard described in A4 was hung on the wall and used throughout A5 to provide visual context for the participant.

4.1.1.3: Procedure

Each participant was asked to provide feedback on the specific high-priority information elements that were identified in A4. Each participant was individually brought to a communal workspace controlled by the research team, which was equipped with the instruments described in the previous section. A short demographic survey was given to record the following data points for later reference in the analysis portion of A5:

- What is your first name?
- What is your age?
- What is your job?
- What is your rank?
- How many years of service do you have?

A PD-based storytelling method was then used to describe the ideal user profile, team member role, common response scenario, and the futuristic storyboard context created in A4. Once this background was set and expectations were established, a series of semi-structured interview questions were asked of each participant:

- What are the different ways to communicate these specific information elements in the common response scenario?
  - Compass
  - Time
  - Radios
  - Messages
  - Gaze
- How would you prefer to have these specific information elements communicated to you in the common response scenario?
  - Compass
  - Time
  - Radios
  - Messages
  - Gaze

The research team then helped the participant record a version of the element as they described it. This sometimes resulted in an aural or haptic modality of communication (e.g., text-to-speech of a message to my radio), but most users preferred to describe visual representations of these specific elements. Myriad drawings were made and digitized for later analysis. When users also described additional elements that were outside the scope of the five-targeted elements for A5, they were recorded as well, but set aside for later use in later Activities. Additionally, when the interviewer perceived from verbal and non-verbal social cues that the participant had more to describe or discuss about a specific element, further open-ended investigatory questions were asked, such as:

- Tell me more about that

- What else?
- What is that?

Responses were recorded again and at the conclusion of digitizing each of the five critical elements in A5, the interviewer asked one final question of each participant:

- Is there anything else you would like to tell me?

Once the interview was finished, the participant returned to his normal workspace and the research team invited the next participant to join them. The digital artifacts of each interview were left on display in the communal workspace so that each participant was able to view previous participants' suggestions. The research team performed a Content Analysis for A5 and then created a final representation standard for each element based on the user-centered feedback that was gathered throughout A5.

The final representation of each information element was verified in a small focus group setting with three first responder participants.

4.1.1.4: Summary

No time limit was enforced during any interview. All additional questions that were asked by any participant were answered immediately. No participant withdrew from any interview. Once all 12 interviews were completed and all information sources were digitized, a brief Content Analysis was performed by the research team in order to:

- Determine which element style was the favorite overall
- Discuss preliminary trends during this ideation exercise
- Determine if additional participants were required for interviews
- Create a recommended standard representation for each information element

4.1.1.5: Results

12 first responder participants were interviewed individually in a semi-structured interview session during A5. As responses were gathered and recorded, the representations of elements became quite

numerous. The following Figure shows several examples of the different UCD information element depictions that were recorded during A5, along with the different instrument mediums that were utilized to create them:

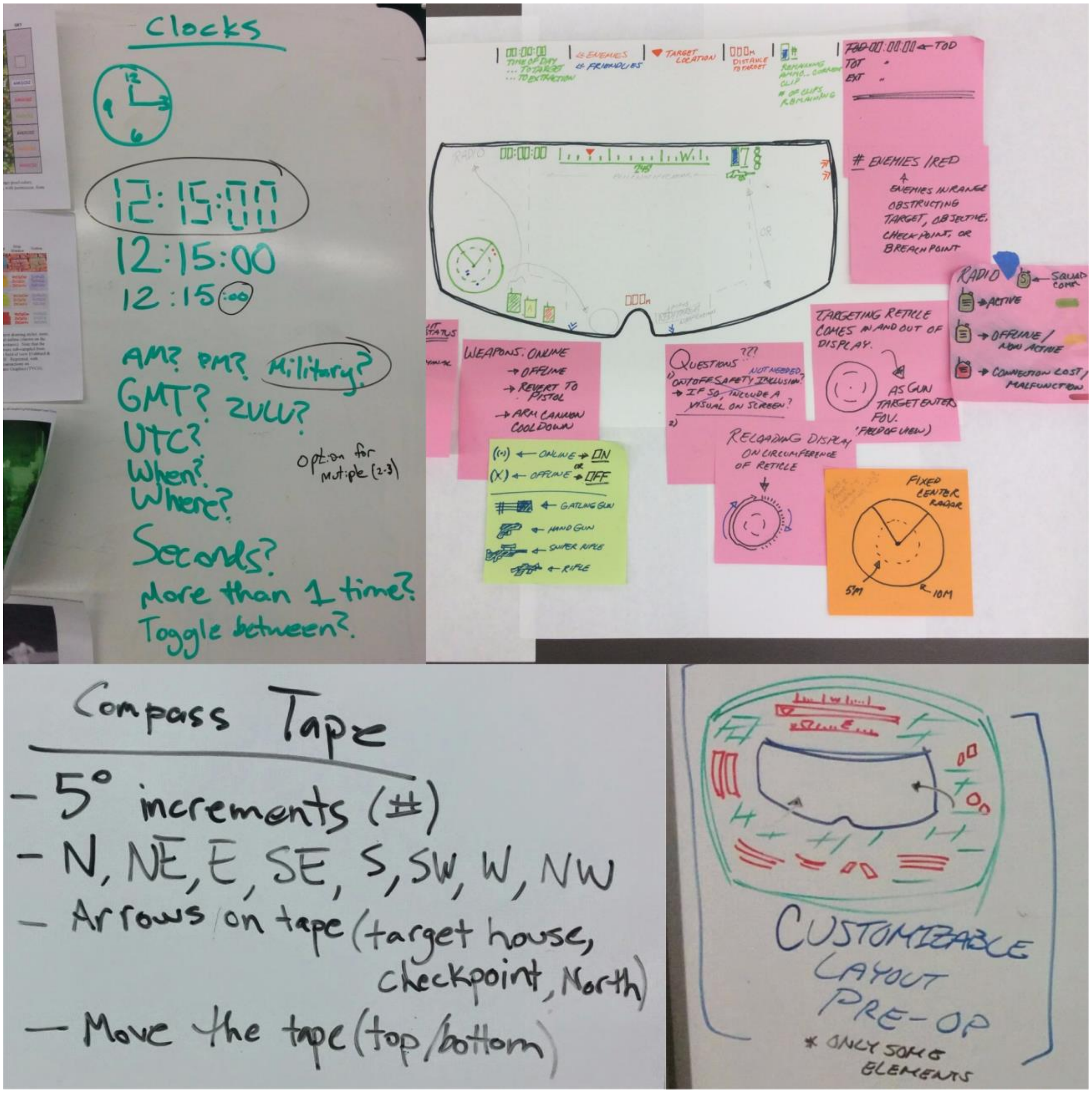


Figure 23: Element Ideation Examples

While the previous Figure shows many different examples of specific elements and feature requests from several participants, it is also useful to show the results of a single in-depth interview discussion with a first responder participant. An example of the final white board concepts from two participants are shown in the following Figure:

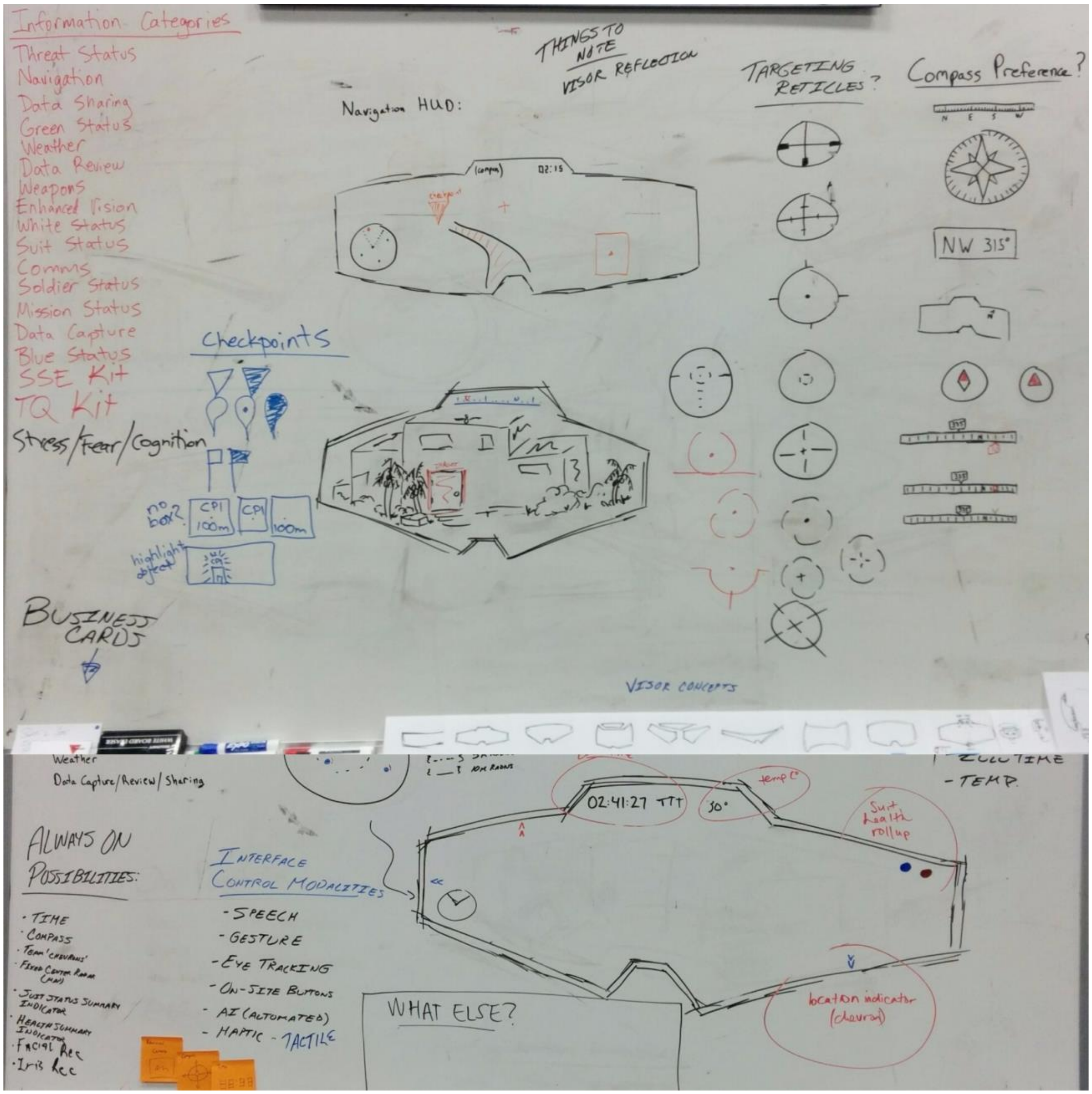


Figure 24: In-depth Interview Notes

Once the standard basic training-based information elements were established, the research team began to start merging these concepts with the common response scenario in its previously established time point organization. The following Figure shows one example of a finalized low-fidelity version of this futuristic first responder display system, created with the assistance of the same graphic artist used in O1:

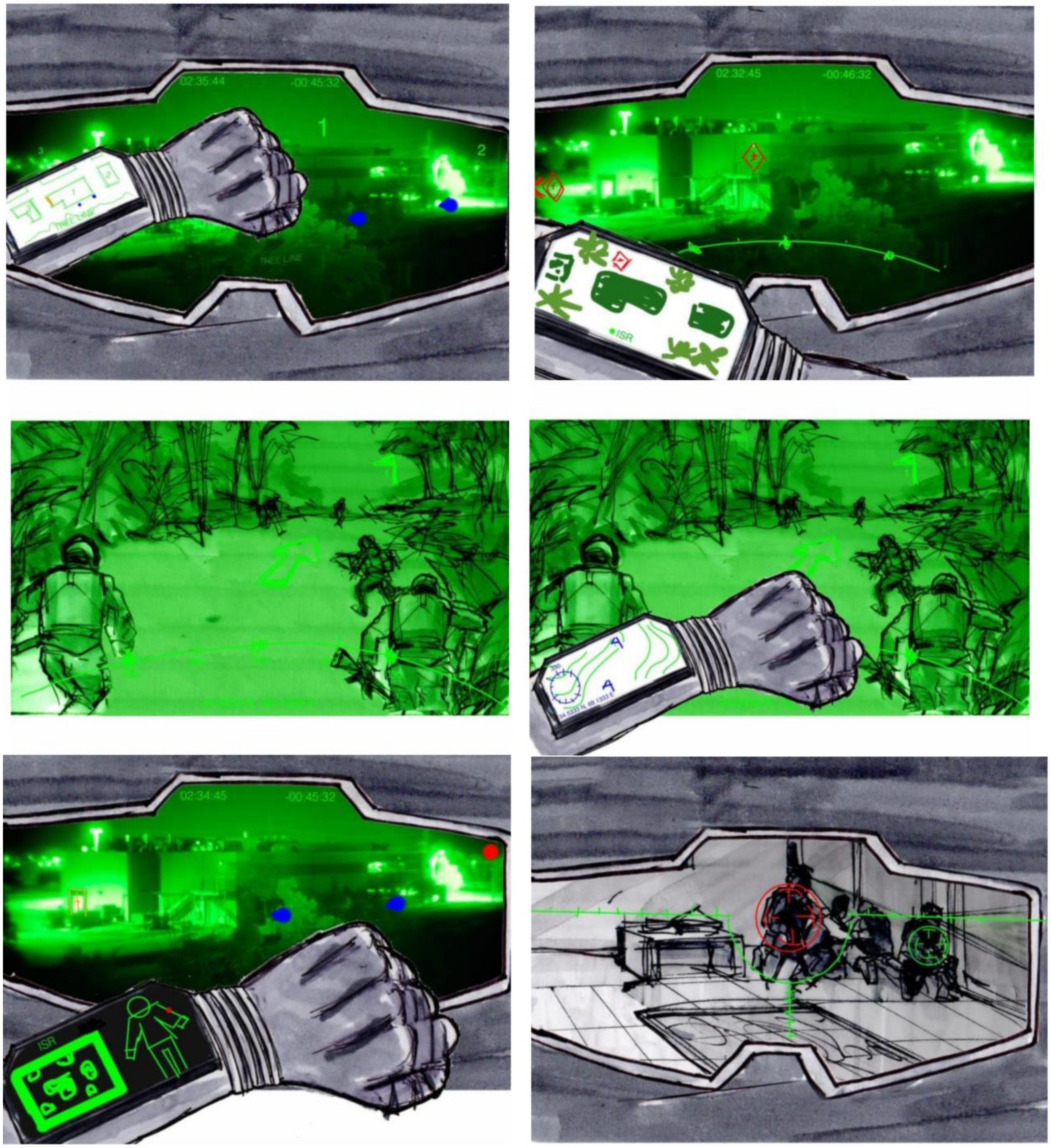


Figure 25: Low-fidelity First Responder Display System Concept

After all 12 interviews were completed, the following data points had been gathered:

- 60 storyboard-matched time point interface designs created; five for each participant
- 50 visual element designs illustrated
- 10 aural element designs described
- 5 haptic element designs described
- A finalized visual "training HUD"

A Content Analysis was performed on the design notes from A5. The results from this methods-based analysis are discussed in the following section of this paper.

4.1.1.6: Discussion

The research team believes that as the ideation process progressed during A5, the element artifacts that were on display in the communal workspace helped later participants to perform an internal analysis of their peers' responses. These existing artifacts were often used to better communicate "their version" of the element (e.g., "I would take that compass and add the numerical degrees to it.").

While all of the information elements were displayed in the communal workspace in some level of fidelity, as the number of visual representations of element data points grew, we believe that later participants were less inclined to suggest alternate modalities of communication. It is postulated that this is partially due to the difficulty in describing these non-visual communication mediums in what became a largely visual workspace. Later experiments in O3 will more formally address this shortcoming.

As information element representations were communicated and created in real-time between each participant and the research team, this allowed for immediate feedback and iteration through the design process. It was easy to create many different representations (e.g., a large quantity of feedback was gathered), in a very short period of time (e.g., low time/effort requirement), thanks to immediate access to first responder participants.

As the research team worked through the analysis of A5, it became more obvious that the actual hardware and software devices that would be used to prototype this 3D MR HWD system would play a critical role in determining the experience of the participant. It was infeasible to create a fully functional high-fidelity user experience in the first prototype of this effort: a low-tech solution was needed instead.

While the collaborative brainstorming session helped to create a large number of information element representations in multiple modalities, the ideation used in A5 also commonly resulted in participant responses that were unrelated to the most critical time point-based information element targets that were the objective of A5. This resulted in many interface designs that were helpful in later Activities, but were out of scope for A5.

As user interfaces were discussed, participants generally had suggestions for user interaction motifs that were also important to them. Gesture control, voice command, and other interaction models were recorded and catalogued for later analysis. Additionally, different levels of overall interaction were discussed by participants, which incorporated single user and multiple-user systems engineering principles. Again, this was out of scope for A5, but provided useful feedback for later Activities.

Customization was perceived as a critical system requirement for all users. Although a standardized set of information elements could be discussed and understood amongst the research team and first responder participants as they related to the training HUD, every participant reiterated their expectation for being able to customize every aspect of the future 3D MR HWD system.

4.2: Activity 6: User Interface Prototypes

While A5 helped to determine the initial design of the most critical information elements of the common response scenario, A6's purpose was to create user interface prototypes that could actually be tested with first responder participants in an experimental setting. The following Figure provides a look at the user-centered design methods employed and the research artifacts constructed during A6:

Figure 26: A6 Overview

4.2.1: Method

4.2.1.1: Participants

A total of 12 participants were used across all interview sessions, six of them were first responders, while the remaining six were simply participants that were familiar with the domain, but had no first responder experience themselves. The non-SME participants' results were maintained separate from that of the first responders so as to not pollute the targeted user population.

4.2.1.2: Instruments

A series of prototype Experiences were iteratively built and evaluated by the research team. Each had different interface, immersion, and interaction characteristics within the experiment. The following seven experiences (i.e., E1, E2, …, E7) were built during A6:

4.2.1.2.1: E1: Projected Elements with Real World

E1 involved projecting the baseline digital element representations that were created in A5 on a blank wall to simulate a possible field of view for the futuristic HWD system. The display hardware utilized a traditional projector with a resolution of 1080p that was shown on a 6.5' screen at a distance of 4' from the participant's head.

4.2.1.2.2: E2: Projected Elements on a Projected Environment v1

E2 added the context of projecting a specific time point on a wall that aligned with the common response scenario to provide a higher fidelity experience of the moment in the scenario that was being communicated to the participant. The display hardware utilized two traditional projectors, each with a resolution of 1080p that was shown on a 6.5' screen at a distance of 4' from the participant's head.

4.2.1.2.3: E3: Mobile Phone with Real World

E3 allowed the user some mobility in being able to traverse a physical space, but at the restriction of a much smaller field of view. This mobile phone-based experience required to user to hold their "HWD" in front of their face in order to view the elements as they were drawn on top of a pass-through video feed of the world in front of them. The display hardware utilized an Android-based smartphone with a resolution of 445p in a 5" size that was shown at a variable distance of less than 2', depending on where the participant held the device.

4.2.1.2.4: E4: Mobile Phone with Low-fidelity VR World

E4 allowed for a low level of user-based customization of the placement of the information elements on a mobile phone-based HWD device with a small keyboard, but at the cost of only displaying a virtual environment. The display hardware utilized an Android-based smartphone with a resolution of 445p in a 5" size that was shown in a VR headset with a 96-degree diagonal FOV.

4.2.1.2.5: E5: VR HWD with High-fidelity VR World

E5 included a completely virtual high-fidelity environment with a consumer-grade development edition VR HWD. A low-level of user-based interaction was allowed here as well. The display hardware utilized a 7" size screen with a resolution of 800p and FOV of 110-degrees diagonal.

4.2.1.2.6: E6: Tethered AR HWD with Low-fidelity VR World

The virtualized world concept from E4 was ported to a pre-release research- and development-based AR-capable HWD. The same low-level user interaction was allowed in this experiment. The display hardware utilized an Android-based device with a binocular optical see-through display with a resolution of 720p with a total FOV of 30-degrees diagonal.

4.2.1.2.7: E7: Tethered AR HWD with Projected Environment v1

The MR-based concept from E2 was ported to a pre-release research and development-based AR-capable HWD. No user interaction was incorporated into this experiment. The display hardware utilized an Android-based device with a binocular optical see-through display with a resolution of 720p with a total FOV of 30-degrees diagonal.

Each prototype in each Experience contained the critical element designs that were identified in A5. The user feedback on each experience was recorded via handwritten notes and used for future development.

4.2.1.3: Procedure

The same process was followed for each experiment. After each Experience was created, it was presented to six participants; three first responders and three non-SMEs to gather feedback on the system. Each participant was individually brought to the same workspace as A5. The same demographic survey from A5 was given here. The same PD-based storytelling method from A5 was used to give the participant the appropriate context of the experiment. Depending on the individual Experience, different instructions were given on how to control the hardware devices (if any user-based control was available). Each participant was then asked the following questions:

- What do you like/dislike about time point 1 in this Experience?
- What do you like/dislike about time point 2 in this Experience?
- What do you like/dislike about time point 3 in this Experience?
- What do you like/dislike about time point 4 in this Experience?
- What do you like/dislike about time point 5 in this Experience?
- What would you change for future Experiences?
- Is there anything else you would like to share with us?

The responses from each interview were recorded and additional follow-up questions were asked when necessary. These results were analyzed by the research team and used to create future Experiences in future Activities.

4.2.1.4: Summary

No time limit was enforced during any interview and no participant withdrew from any interview. The user feedback gathered after each Experience was used by the research team to design the next Experience. A Context Analysis was performed on the written responses in order to:

- Discover preliminary trends
- Determine which areas of the Experience required further research and exploration
- Decide how many additional participant interviews and Experiences would be required to understand the user interface enough to progress in the subsequent design and prototyping efforts

Researchers attempted to keep each participant interview as short as possible throughout A6, but allowed the participant to speak as freely as they wished.

4.2.1.5: Results

The seven Experiences of this Activity resulted in many different research artifacts. Additional detail on each of the Experiences will be described in this section, including relevant example figures where appropriate.

4.2.1.5.1: E1: Projected Elements with Real World

E1 included a refined digital version of the information elements that were developed in A5. These elements were placed within a presentation file format that could be displayed via a projector on a wall within a workspace. This allowed the research team, first responder participants, and supporting personnel to all experience what a futuristic MR HWD system might feel like. Feedback was gathered from participants and used to inform E2. An example slide from this presentation is shown in the following Figure:

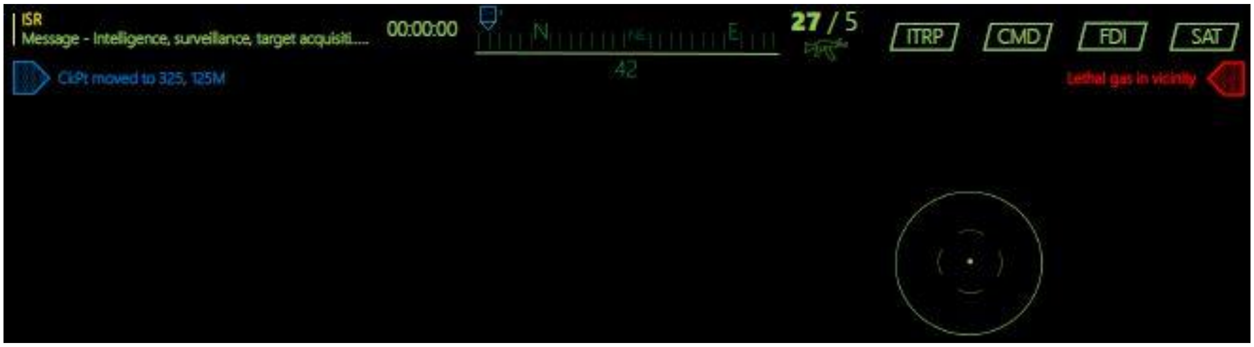


Figure 27: Experience 1: Projected elements on a blank environment

E1 was a good start, but the experiment participants all indicated that increased user interface fidelity was needed in future experiment iterations to provide more realistic feedback.

4.2.1.5.2: E2: Projected Elements on a Projected Environment v1

E2 included the same presentation file as E1, but included additional time point-sensitive context during this experiment. A second projector system was added to E2 to display these time points, with both projectors being overlaid on each other. Feedback was gathered from participants and used to inform E3. An example combined slide time point from this experiment is shown in the following Figure:

Figure 28: Experience 2: Projected elements on a projected environment

E2 was a good next step in this effort, but all participants wanted to have more user interface fidelity in future iterations.

4.2.1.5.3: E3: Mobile Phone with Real World

E3 was a first attempt at an AR-based HWD system. Using a custom mobile phone application, the research team was able to view representations of 2D information elements placed on a 2D viewing device (one camera and one display screen). User-based feedback from this experiment was recorded and used to quickly iterate to E4. A screen shot of this application in action is shown in the following Figure:

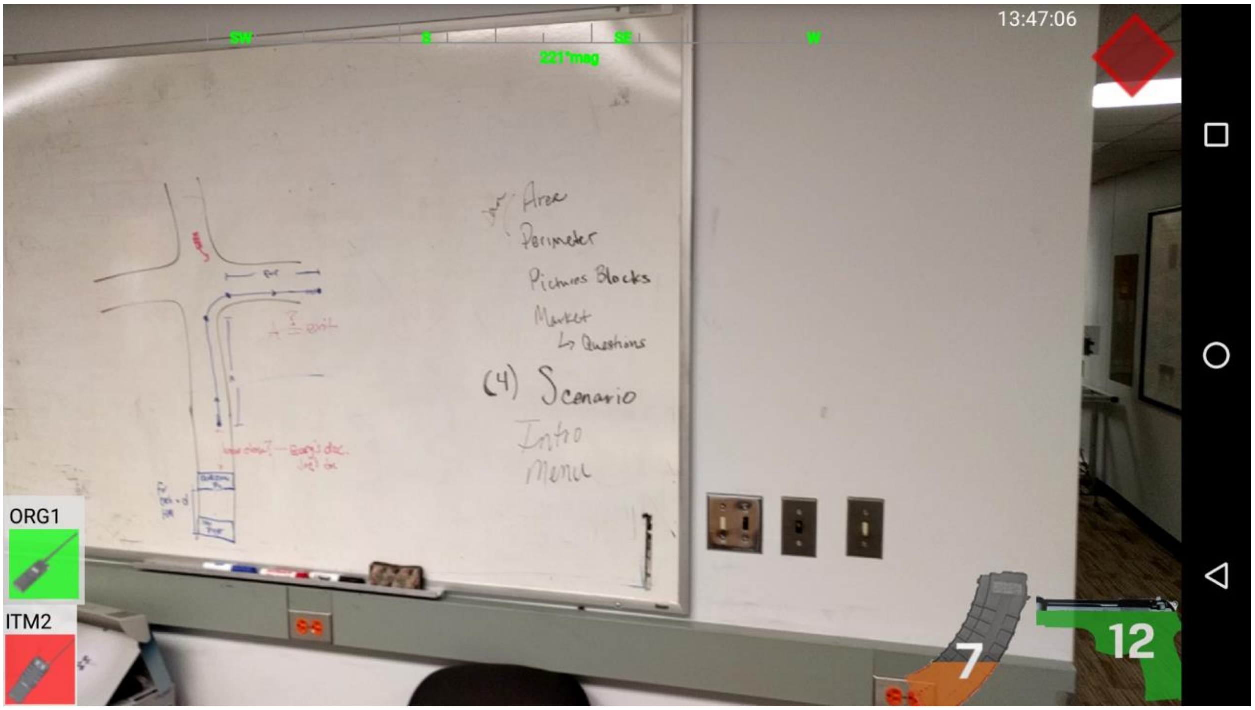


Figure 29: Experience 3: Pass-through camera mobile application

E3 provided a much smaller FOV, but was the first Experience that was mobile. All participants communicated their desire to have more user interface fidelity in future iterations.

4.2.1.5.4: E4: Mobile Phone with Low-fidelity VR World

A VR-based mobile phone application was used in this experiment. Although E4 was a completely virtualized environment, it did include the first attempt at including a user interaction in the experiment. Although crude, this application did allow for users to customize where some information elements were located on their screen. The following Figure shows a screen shot from this application as it was operated:

Figure 30: VR-based mobile HWD

The following Figure displays the list of available control commands provided to the user during E4:

For Both:

Left/Right Arrows to rotate view
Spacebar to shoot
Q, W, E, R to speak into different radio channels

To manipulate location of display elements:
Hold down:

- **1** for Radio
- **2** for Notification window
- **3** for Clock
- **4** for Compass Tape
- **5** for Weapon Status
- **6** for AR Overlay

And:

- **Arrow Keys** to move element
- **+,-** to toggle element on and off

Press X to reset location of elements

For Full Simulation:

V to begin background video as environmental alternative

Figure 31: E4 User Interaction Controls

And the following Figure shows a third-person perspective of E4 in operation during an experiment:

Figure 32: E4 Third-person Perspective

Because participants were not able to see the control devices employed in E4, due to the fully-occluded HWD, experimentation was abandoned by the research team and E5 was developed to replace E4. Researchers were given enough time during the development of this experiment to memorize the operation of E4, but such an expectation was unrealistic for first responder participants.

4.2.1.5.5: E5: VR HWD with High-fidelity VR World

A consumer-grade development version of a popular tethered VR HWD was used in this experiement. E5 simplified some of the user interaction in E4, but provided a much larger FOV at a much shorter level of latency than the phone-based applications in prior experiments. An increased fidelity in terms of user interface elements was also explored during E5. The user was also allowed to move themselves around in the virtual environment. Learnings from gathered participant feedback in E5 led to the development of E6.

4.2.1.5.6: E6: Tethered AR HWD with Low-fidelity VR World

The research team acquired a beta version of a professional-grade AR-capable HWD during A6. The virtualized world from E4 was ported to this highly specialized HWD and the same level of user interaction from E4 was available for testing in E6. Although a VR world was used on this HWD, which generally occluded the real-world environment, the participants still had enough peripheral vision that they could manipulate a secondary control device and customize their screen. These initial screen preferences were saved and digitized for future reference. The feedback from participants was gathered and a Content Analysis was performed on the results of all previous Experiences.

4.2.1.5.7: E7: Tethered AR HWD with Projected Environment v1

This final Experience within A6 served as the apex-level event for this portion of the research effort. It was the final physical demonstration of the futuristic HWD for first responders that was developed during O2 in order to highlight the system designs for the user interface. Based on previous participant feedback from E1-E6, the research team used the AR-capable HWD from E6, combined with the projected environments from E2 (which provided the realistic context) and the information elements from E1 (in digitized 2D designs) to create E7. The following Figure shows two perspectives from this experiment:

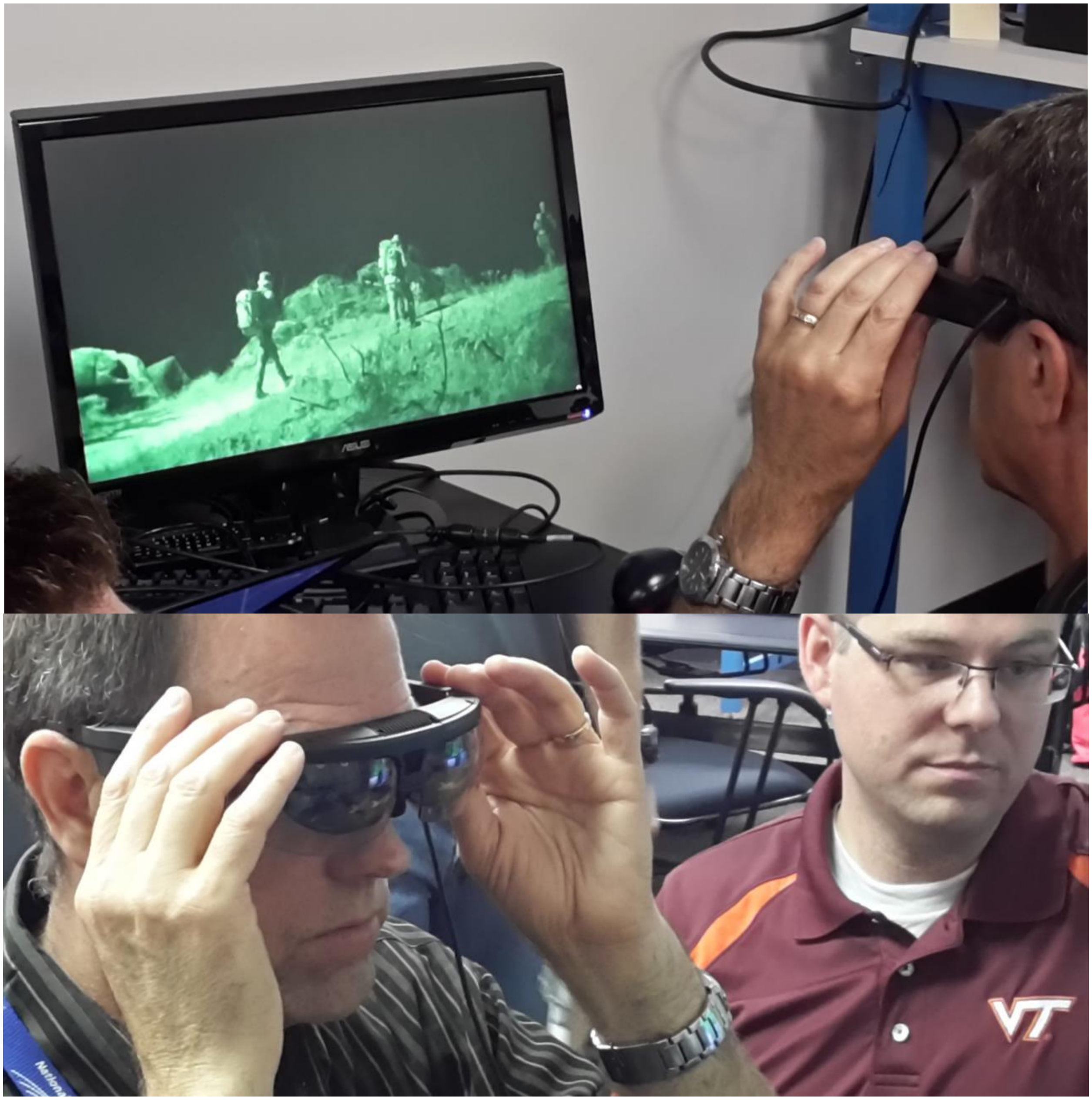


Figure 33: Over the Shoulder and Participant View of E7 Experiment Session

E7 allowed for minimal user interaction in order for the research team to better control the user interface prototype elements included in the experiment. Feedback from E7 was gathered and analyzed. The discussion below details the lessons learned throughout A6.

4.2.1.6: Discussion

Quick and rapid iterations in prototyping allowed for seven Experiences to be explored during A6. Although each was short in its duration, there were many notes and research artifacts for the research team to analyze. The semi-structured participant feedback from these Experiences led to many important decision points during this portion of the research effort. Each Experience had its own unique benefits and drawbacks. Individual lessons learned during each Experience will be shared in the following sub-sections. A final discussion section for A6 will then be described overall.

4.2.1.6.1: E1: Projected Elements with Real World

Beginning with a very low-fidelity user interface and user interaction experiment, the research team could quickly gather initial impressions on how the information elements that were developed in A5 would actually be displayed in a futuristic 3D MR HWD system for first responders. This allowed for final changes to the designs of information elements (colors, sizes, shapes, etc.) in a rapidly iterative manner. This also standardized the training HUD described in previous Activities. What was completely lacking from E1 was mobility in the Experience and any sense of a realistic scenario to provide context to the first responder.

4.2.1.6.2: E2: Projected Elements on a Projected Environment v1

E2 added a realistic time point-based environmental context to the testing scenario. This was helpful for the PD-based storytelling method goal of placing the participant in the proper mental situation of the moment that was being discussed.

Users continued to remark that they wanted more user interface fidelity in future experiments. The elements themselves were visually well-designed and displayed within a good context at this stage of A6, so little was remarked in relation to any element changes, such as apply to the size, shape, color, etc. Researchers then decided that a mobile version of E2 should be created in order to view the individual elements in the contexts of varying real-world background colors and patterns. This would further progress toward the goal of a futuristic 3D MR HWD system that first responders could actually use in the field.

4.2.1.6.3: E3: Mobile Phone with Real World

E3 included a simple Android-based pass-through camera mobile phone application. This application allowed for 2D information elements to be placed on a 2D screen that showed a video version of the real world in front of the phone. The user was required to hold the mobile phone with their own hands in this experiment. This gave a more realistic sense of the environment and brought to light how some information element designs would not be feasible in some scenarios, due largely to color choices of elements. It became especially obvious that the characteristics (color, brightness, etc.) of real-world objects often would occlude what was displayed on the HWD, so a more controlled testing scenario was necessary in future experiments in order to keep the user focused on gathering feedback on the user interface prototypes themselves. A color-limited night-only mode of operation for these experiments would later become the time of day that was preferred for experimental purposes in order to decrease the issues that arise from daytime HWD HUD operation, similarly to what previous researchers in this domain have indicated for many years (J. L. Gabbard, Swan, Zedlitz, & Winchester, 2010; Joseph L. Gabbard, 1997).

The largest drawbacks from E3 were the limited FOV (determined by the resolution and screen size of the device) and the requirement of the user to hold the device in order to view it. It is also important to note that a 2D screen and 2D objects were used because no quickly-deployed 3D option was available to the research team at the time E3 was developed. These shortcomings will be addressed in later iterations and Experiences. Users continued to mention their desire for more user interface fidelity in future experiments.

4.2.1.6.4: E4: Mobile phone with Low-fidelity VR World

E3 marked the first experiment that utilized a visually-occluded HWD. The physical hardware used in E4 was the same as E3, but included a virtually generated environment in place of a camera pass-through video mode of operation. In this VR-based experiment, the first attempt at creating 3D user interface objects and user interaction motifs was accomplished. Additionally, the VR HWD could be placed within a holding device so the user's hands were free to operate secondary control devices.

The largest drawback from E3 was the occlusion of the real-world environment. Although the research team had enough practice with secondary user interaction devices during development that they could control them without visual reference of the controllers themselves throughout an interaction session, the first responder participants were not able to operate the controls without seeing the real-world device. Although experiments were performed, because the targeted users were unable to complete the task, the semi-structured interview was ineffectual in gathering useful feedback on the user interface and user interaction system beyond the truth that no user was able to customize their HUD. A complex interaction system, therefore, requires an accurate virtual analog system within a VR environment experimentation scenario.

4.2.1.6.5: E5: VR HWD with High-fidelity VR World

E5 moved to a higher resolution consumer-level development version of a VR HWD. This also included a larger FOV with additional user interaction devices. With this additional functionality, a higher-fidelity virtual environment was created in which the user was allowed to move themselves around. The same HUD customization in previous experiments was also present in E5. Users generally enjoyed this more in-depth and immersive user interface and user interaction Experience, but called for an MR version where the real world was not occluded.

While most users enjoyed this experience overall, the user interaction level was very basic, involving moving the character around the virtualized world. HUD customization was available to perform, but the research team was unable to provide an analogous virtual representation of the control devices that were available to them within the VR world. This, in turn, meant that a user was still required to memorize how to operate the user interaction device. This proved less effective than hoped when tested with first responder participants due to the training time required to learn the interaction system before a user would become proficient in its operation. Users again repeated their desire for more user interface fidelity and capability in future experimental settings.

About 40% of all users who tested this VR system (including the research team and other co-workers) suffered from simulation sickness; thereby discouraging the use of this fully occluded technology in future Experiences. Every attempt was made during the development of this experiment to decrease the nauseous feeling of simulation sickness, but it was never removed to the degree that everyone could participate in E5 without motion sickness. While there were many different iterations of this application within E5, only the final version was tested with the first responder participant pool. The level to which simulation sickness could potentially negatively impact the perceived and actual functionality of a futuristic 3D MR HWD system designed specifically for first responders meant that future Experiences would need to rely on MR-based HWDs whenever possible.

Finally, a large amount of resources were dedicated to specialized programming that could only be used in the VR HWD for E5. While this provided a more immersive virtual Experience to participants, this also meant that a large portion of the effort expended to make E5 a worthwhile experiment could not be used in future MR-based HWDs.

4.2.1.6.6: E6: Tethered AR HWD with Low-fidelity VR World

Starting with the code base from E4, the research team ported that concept to the E6 testing scenario. The greatest change that occurred by this point in the research effort was a cutting edge professional grade AR-capable HWD that was acquired by the research team. Essentially a beta release of an Android-based mobile device packed into the form factor of a mobile HWD, this headset would allow the mobile applications of previous experiments to be used in a much more convenient, hands-free HWD.

Because the virtualized world of E4 was being reused in E6, the AR feature of this new headset was not maximized during experiment. Although the VR environment was being projected through the display of the HWD, it did not fully occlude the real-world environment of the participant pool. This allowed for a relatively fast iteration of the E4 code base to the new HWD where the first responder participants were finally allowed to fully customize their HUD as they saw fit.

There were many drawbacks to the overall E6 scenario. The largest complaint from the first responder participants was the smaller FOV with the AR-capable HWD. While this was an understandable tradeoff to the research team, the first responders responded very negatively to the reduced FOV. This smaller FOV also meant that the movement of the information elements was highly restricted; all users essentially placed the elements at the bottom of the HWD’s visual area. Many participants also remarked that they would decrease the number of elements displayed because of this smaller operational area.

Because this was a beta release product, there were many software and hardware issues to contend with. Overheating, system crashes, and/or user interaction device inconsistencies occurred during nearly every data-gathering semi-structured interview session with first responder participants. Regardless of the precautions taken by the research team, this specific HWD was quite unreliable in its pre-public release form. This showed the research team that they needed to control as many of the variables of E6 as were possible in order for future experiments to be useful to the overall progress of the entire research effort.

4.2.1.6.7: E7: Tethered AR HWD with Projected Environment v1

E7 was the final experiment within A6 and served as the cumulative exemplar of what a futuristic 3D MR HWD system for first responders could be. Due to the software and hardware complexities and inconsistencies of previous Experiences, E7 was designed to be a high-fidelity visual user interface experiment, but without user interaction capability. In technical terms, the HWD of E6 was reused in this experiment, which displayed a highly customized set of information elements developed from E1. At the same time, the time point scenario examples of E2 were projected on a large wall in front of each participant. This gave an immersive and realistic feel to the testing scenario, but with the addition of a much more realistic HWD than had been utilized previously. In order to sync all of the display systems in E7, all display surfaces were tethered and controlled by a series of three laptop computers.

Each participant reiterated their desire to have a larger FOV, higher fidelity user interfaces, and user interaction capabilities in future experiments. Some software and hardware bugs persisted in this highly-controlled experiment, but these were dramatically reduced compared to E6. Additionally, the research team had been using the HWD of E7 for a few weeks at this point in the research effort, so working around the experiment’s glitches had become second-nature.

4.2.2: Additional Discussion

As participant feedback was gathered, the research team attempted to gradually increase the level of user interface fidelity. As will be continually referenced throughout this paper, every first responder participant always requested more user interface and interaction fidelity during every semi-structured interview session. When determined to be easy to add, interactions were included in various experiments (i.e., E4, E5, E6), but often proved distracting to the overall purpose of the experiment in relation to increasing the user interface fidelity of the system.

Interestingly, from a post hoc analysis perspective, only E3 allowed for truly non-tethered experimentation. This was another complaint that users consistently provided during every semi-structured interview session that involved a tethered experience (i.e., E1-E2, E4-E7). To emulate a futuristic 3D MR HWD system for first responders, it was important that future Experiences include non-tethered interface and interaction devices.

In order to rapidly prototype, no Experience could be ranked as truly high-fidelity at this phase of the research. Rapid iteration required a small scope and a low level of resource dedication for this effort. While some Experiences (E5, E6, E7) relied on previous work, when it was available, each experiment had to be designed, developed, and executed in a matter of days. The largest constraint during A6 was the limited time to complete this portion of the overall research and gather data-based output artifacts for use in later design and rapid prototype iterations.

By the end of A6, the research team and first responder participants were comfortable with the design and visual execution of the basic training HUD with the test HWD technologies. From basic hardware projector tests to cutting-edge beta-release AR HWDs, a dynamic range of experimentation testbeds were employed throughout A6. A firm understanding was achieved by the end of A6 of the common response scenario as it relates to a single assaulter acting within the futuristic 3D MR HWD environment. The research team now felt qualified to explore additional team member roles and increase the user interface and interaction fidelity of the common response scenario.

4.3: Activity 7: Storyboard 2.0

While A6 helped to create many different user interface Experiences that could be tested with first responder participants during fast, iterative evaluations, A7 was designed to add breadth to the user interface scenario and prepare for an interactive user interface prototype system in O3. The following Figure shows the UCD design methods employed during A7, along with the research artifact outputs:

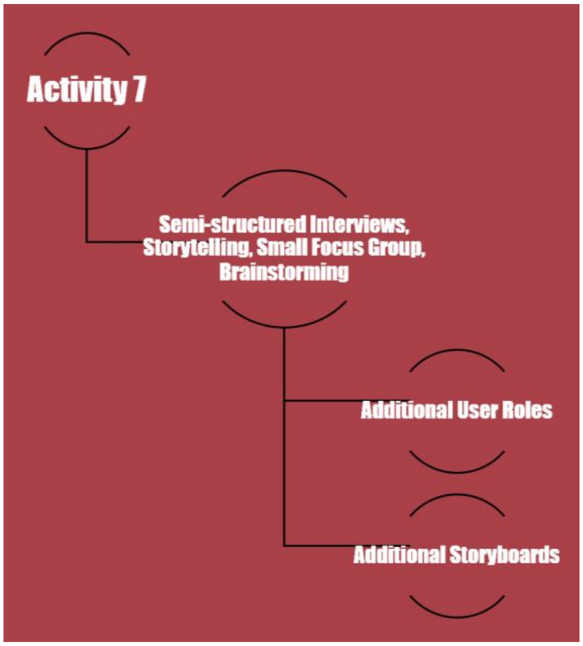


Figure 34: A7 Overview

4.3.1: Method

4.3.1.1: Participants

Six first responder participants were individually interviewed in a semi-structured fashion throughout A7. Two of these participants became SME representatives for A7 and were integrated into the research team for approximately two work weeks. A final small focus group was held with four first responders to critique Storyboard 2.0.

4.3.1.2: Instruments

White board sketches and other physical and digital creation devices were used as instruments to build analog and digital artwork. Additional notes were taken during interview sessions, digitized, and later analyzed by the research team.

4.3.1.3: Procedure

Each participant was asked to provide more detailed information about the common response scenario. Initially, the participants were questioned as to which two team member roles from O1 might benefit the most from a futuristic 3D MR HWD system for first responders. All participants agreed that a command and paramedic team member perspective would greatly benefit from such a capability.

With the command and paramedic team member perspectives in mind, the five time point organization and PD-based storytelling methods employed in previous Activities were used again here to place each participant in the common response scenario. Being that these additional roles had generally been ignored by the research team prior to A7, a brief experimental version of O1 was executed in order to gather the necessary information on these additional roles and to rapidly design and prototype the command and paramedic team member perspectives. The following questions were asked of each first responder participant, adapted from the same types of questions asked throughout O1:

- At time point X (where X represents 1-5)
  - Which kinds of information do you use today to do your job?
  - What pieces of information do you utilize during this time point?
  - How are these information elements communicated to you today?
  - If anything were possible, what kinds of information would you like to have access to?
  - What pieces of information would you like to utilize during this scenario?
  - How would you like to be notified of an information element?
  - What is the best way to communicate an information element to you?

- While you are doing your job, what information elements are the most important for you to have at this time point?
- Of these elements, what is the priority of information to communicate to you?

The research team, now with the benefit of having worked in the first responder domain for the previous year of this effort, was able to much more quickly gather the general information required from these semi-structured interviews. When necessary, additional follow-up questions were asked and discussion related to the common response scenario was recorded via handwritten notes. When the initial six interviews were finished, researchers had enough information to begin creating prototype designs and research artifacts for review and experimentation with first responders. Two first responders volunteered to integrate into the research team, represent the voice of the SMEs, and rapidly iterate through myriad design sessions to create a newly updated version of the storyboard originally developed during O1. During this period of storyboard development, the most critical information elements identified during each time point were standardized and vetted with first responders, as was also done during A4. A final small focus group session was held with four first responders to ensure the accuracy of Storyboard 2.0.

4.3.1.4: Summary

Notes from the interviews were recorded by the research team, digitized, and a Content Analysis was performed after each participant's responses. No time limit was enforced for A7 and the first six interviews lasted approximately one hour. The iterative design work performed with the two SME representatives within the research team lasted approximately two work weeks. Hundreds of hours of design and prototype work were performed during A7, spanning approximately three work weeks of effort. A final focus group was held to verify Storyboard 2.0.

4.3.1.5: Results

One consistent complaint from first responders during this entire research effort that has not been detailed previously had been the focus on only one membership role in the overall team. This was initially limited by design in order to reduce the scope of the effort and ensure that the concerns of the most endangered team member role were addressed first. It was also necessary to limit data input from first responders because the research team had no prior experience in this domain. However, by the end of A6, it was clear that the research team was competent in the first responder domain and capable of exploring additional team roles to add breadth to the effort. These two roles were researched in-depth and a new storyboard was created, similar to the one discussed in O1 of this report. The revised storyboard included a much greater level of detail for the common response scenario, along with the three chosen perspectives, which would be used for this and all future Activities.

One example of an ideation session from one time point of the medic perspective is shown in the following figures. The first shows the output of one interview in terms of the priority of information that needed to be communicated to the first responder during one time point of the common response scenario:

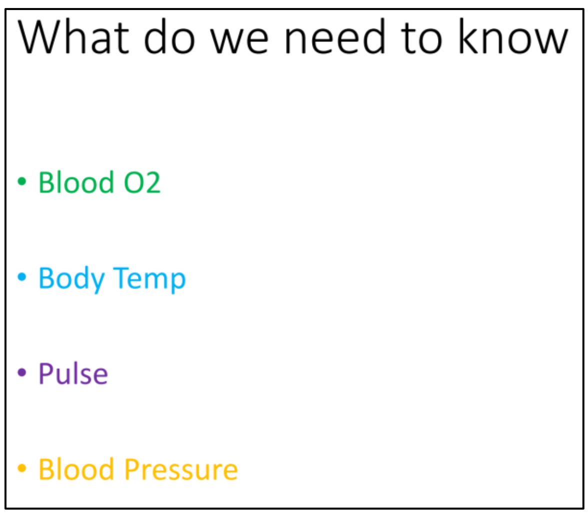


Figure 35: Identified Priorities of Information for Paramedic Perspective

The next Figure shows an ideation design and prototype session with one participant that helped identify what specific information elements would be displayed and how those would be communicated, similarly to the process performed during O1:

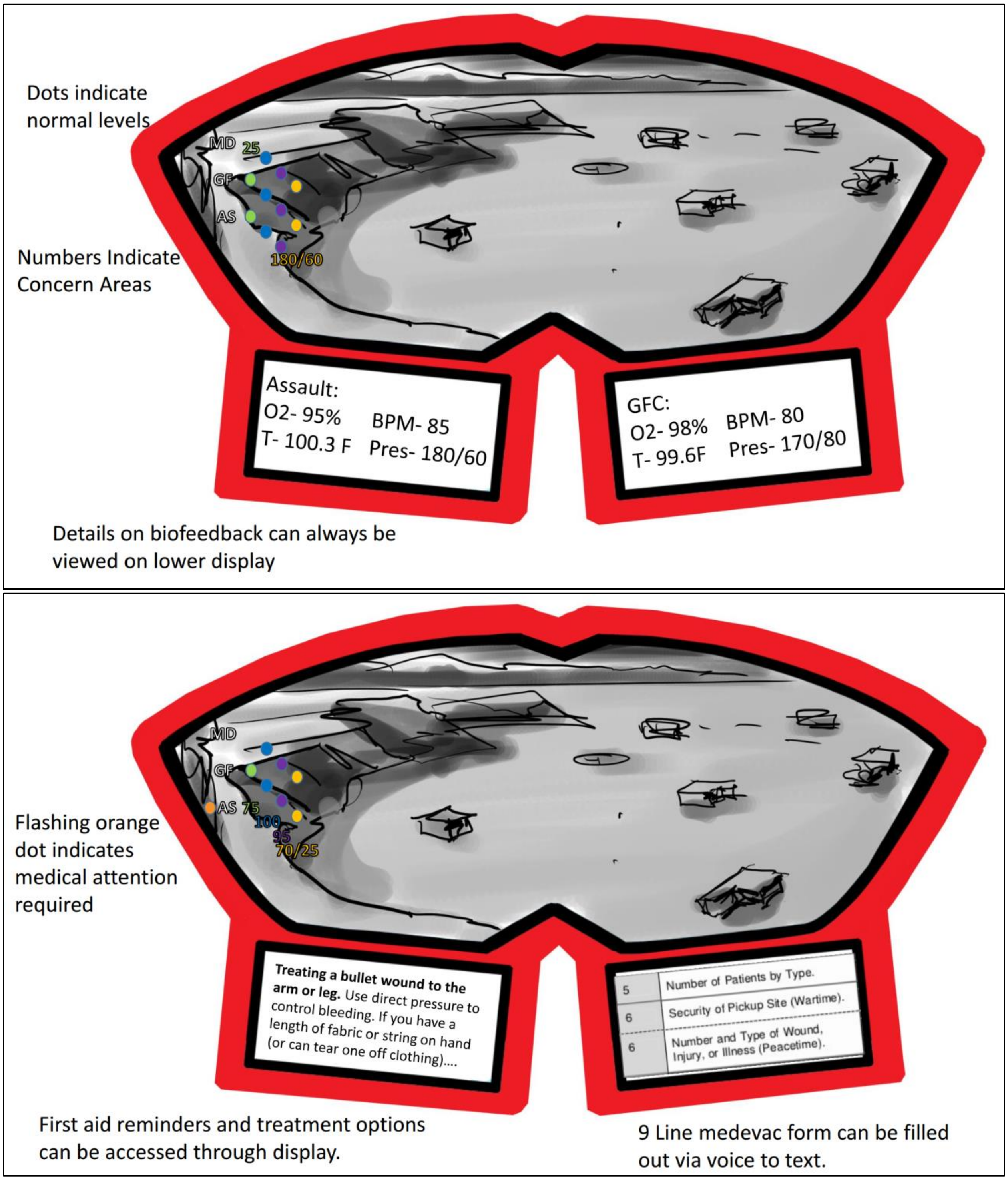


Figure 36: Ideation for Medic Perspective

The following Figure shows a picture of one of the research team design sessions that was held during the storyboard creation, including two artists, a UCD researcher, and a first responder SME representative:

Figure 37: Storyboard Creation Session

A final storyboard was developed by the research team and verified by first responder participants in a small focus group. Due to the complexity and length of this document – a 14-page, 4K resolution artifact – the storyboard is not included as an in-text citation in this chapter. See Appendix A for the revised Storyboard in its entirety.

4.3.1.6: Discussion

Because of the addition of team-based perspectives, it was necessary to develop many new information element representations. A simplified and expedited version of the same information gathering processes followed in O1 was performed during A7. New elements were prioritized, designed, and vetted with first responder participants. This could be performed in a rapid and iterative fashion due to the domain knowledge of the research team by this stage of the overall effort, in addition to the makeup of the research team itself. First responder SMEs, UCD researchers, artists, and graphic designers were all co-located in the same physical workspace and were able to work together every day to progress the overall research effort. This rapidly iterative design and prototyping could not have been performed as easily in a telecommuting-type environment.

While some small issues were encountered during A7 (e.g., SMEs not always agreeing on what to name a specific element), no large methodological interruptions occurred. The PD-based storytelling method, combined with semi-structured interviews, small focus groups, and brainstorming sessions were very effective in developing the storyboard during A7.

It is important to note that while this more in-depth storyboard was developed, the main focus of the research effort remained on the most endangered team member: the first person to answer the call of the common response scenario. However, the first responder participants reacted very positively to the revised storyboard and were very clearly engaged in its creation; they appreciated the work that went into creating it and were happy to spend time developing it alongside the research team.

5: Objective 3 (O3): User Interactions

For the benefit of the reader, the following Figure communicates the three overall Objectives of this effort. See previous chapters for more details related to this graphic:

Figure 38: Abbreviated Research Effort Overview

Along the course of interface development, this work will continue to examine human factors research questions relevant to each objective. The following chapter describes O3 of the effort, which encompasses a description of rapid design and prototype user interaction-based experimentation. O3 answers RQ5 and RQ6. The stated RQs of O3 are as follows:

- RQ5: How do first responders desire to interact with critical information?
- RQ6: How does the context in which a user experiences a prototype effect the interaction feedback they provide?

The following Figure depicts an overview of the planned methods that will be used during O3, along with the resulting research artifacts:

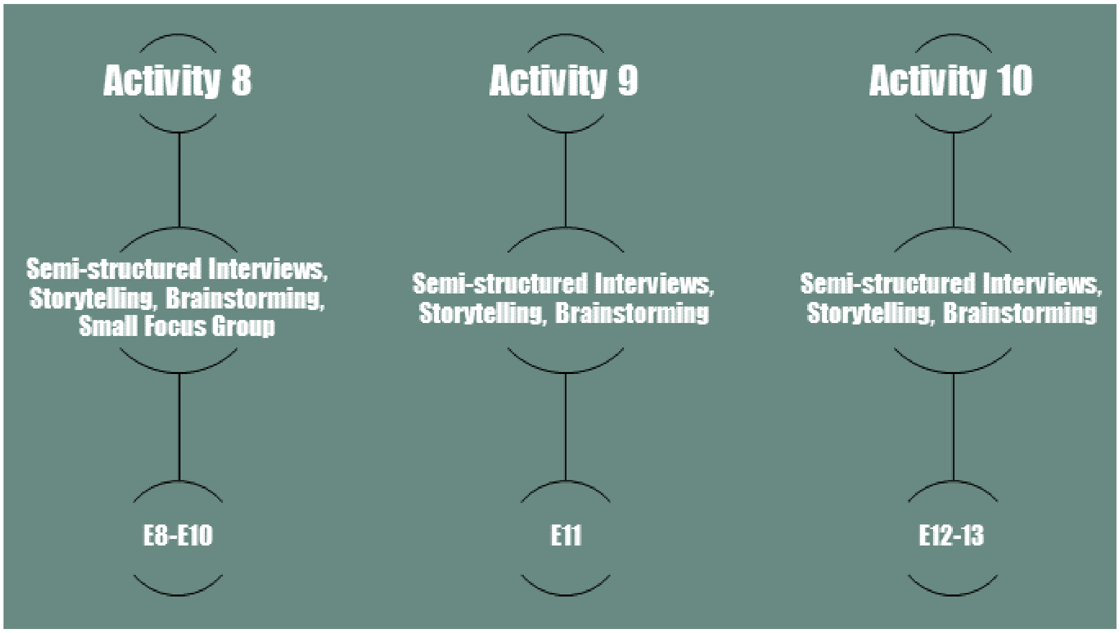


Figure 39: Objective 3 Overview

Like the discussion in the previous chapter, this chapter includes three unique Activities (i.e., A8, A9, A10) that comprise the research work of O3. With the completion of the initial information requirements gathering of O1 and user interface design and prototype iterations in O2, the research team performed a final series of iterative design and prototype phases during which 3D MR HWD system user interactions were explored. O3 was informed by the results of O1 and O2. A post-mortem analysis of the methods and tools utilized during O3 will be described later for this portion of the research effort.

5.1: Activity 8: User Interaction Prototype

The original purpose of A8 was to create a high-fidelity user interaction prototype that was based on the lessons learned from previous Activities. It would incorporate the consistently requested high-fidelity level of operational user interaction between the participant and the HWD system. The formal addition of user interaction design and prototyping in A8 marked an important milestone in terms of the research team's response to the participants' continued request for greater levels of interface and interaction capabilities in the prototype system. The following Figure shows the UCD-based methods employed by the research team during A8, along with the research artifact outputs created:

Figure 40: A8 Overview

5.1.1: Method

5.1.1.1: Participants

18 total participants were interviewed individually in a semi-structured interview session during A8. Nine of the participants were from the ideal first responder user profile while the remaining participants were non-SME users that could provide more technical feedback on the overall user interface execution and interaction motifs of the experiments.

5.1.1.2: Instruments

Handwritten notes, white board concepts, and digital system engineering files served as the instruments used here. All white board research artifacts were digitized via photograph and stored for later analysis by the research team. Three different Experiences were developed during A8. These Experiences are described in more detail in the following sections:

5.1.1.2.1: E8: Tethered AR HWD with High-fidelity VR World

E8 utilized the same AR-capable HWD from E6 and E7, but with a high-fidelity VR environment projected onto the headset to provide additional situational awareness. E8 also had several user interaction control devices integrated into the Experience, including the following devices:

- Small Bluetooth keyboard
- Finger-operated mouse
- Gesture-tracking haptic response armband
- Touch pad mouse
- Touch-sensitive glove
- Bluetooth direction pad
- Earphones (for audio feedback to the user)

This resulted in an experimental testbed with visual interface similarities to what was executed in E6, but with a suite of third-party consumer interaction devices that were designed to allow the user control of the HWD system. The display hardware utilized an Android-based device with a binocular optical see-through display with a resolution of 720p with a total FOV of 30-degrees diagonal.

5.1.1.2.2: E9: Tethered AR HWD with Real World v1

E9 was a simplified version of E8 with the removal of the high-fidelity VR world used in that Experience. The same user interaction options existed from E8. The display hardware utilized an Android-based device with a binocular optical see-through display with a resolution of 720p with a total FOV of 30-degrees diagonal.

5.1.1.2.3: E10: Large FOV AR HWD with Real World

E10 was a further simplified version of E9 with the removal of all user interaction devices from this experiment. It was meant to showcase a large FOV AR HWD. The hardware utilized a binocular optical see-through display with resolution of 1200p with a total FOV of 120-degrees diagonal.

5.1.1.3: Procedure

Three Experiences were developed during A8. Each release candidate version of each Experience was tested with first responder and technical experts. The pre-release development versions of each iterative prototype were only tested internally with the research team. The same five time points of the common response scenario were explored and served to mentally structure each experiment. PD-based storytelling methods added further context to the situational awareness of the participant. The command-level team member profile was the targeted user for E8 and E9, while the assault-level team member profile was the targeted user for E10. Users were given a brief verbal tutorial about the interaction systems included in the experiment, including voice and gesture capabilities, along with physical buttons that would change the information displayed on the HWD. In all semi-structured interviews with end-users for E8 and E9, the following questions were asked, similar to the questions asked during A6 of the research effort:

- At all time points:
    - Perform a voice command to change your HUD.
    - Perform a gesture command to change your HUD.
    - Press a physical button to change your HUD.
    - What do you like/dislike about this Experience?
- What would you change for future Experiences?
- Is there anything else you would like to share with us?

The semi-structured interview questions for E10 were as follows:

- At all time points:
    - What do you like/dislike about this Experience?
    - What would you change for future Experiences?
    - Is there anything else you would like to share with us?

The results from each test were recorded and later analyzed by the research team. A final small focus group session was held with two first responders and the research team to summarize the findings of A8.

5.1.1.4: Summary

Notes from each of the three Experiences were digitized and a post hoc Content Analysis was performed by the research team.

5.1.1.5: Results

The end of A7 marked the decision point of the research team to formally add user-based controls and interaction motifs to the first responder testbed system that would be implemented in O3. Because previous attempts at integrating user-based controls had not been the focus of prior research experiments and had generally failed to provide useful user-centered feedback, a more formal design of the concept was needed. An example of one ideation session for a control system is shown in Appendix B.

Several control systems were also designed conceptually and were physically implemented in later Activities. In addition to the more in-depth storyboard scenarios from A7, increasing the fidelity of the system was important to give more realism to the experimentation. User-based interaction was designed and implemented in several different modalities:

- Voice command input to system
- Gesture input to system
- Physical controller input to system (via various controller buttons)
- Haptic feedback to user
- Voice feedback to user
- Visual feedback to user

Each participant was asked to give general feedback on what they did and did not like about each Experience. The three Experiences of A8 resulted in different research artifacts. Additional detail on each of the Experiences will be described in this section, including relevant example Figures where appropriate.

5.1.1.5.1: E8: Tethered AR HWD with High-fidelity VR World

E8 utilized the AR-capable HWD from E7 within the context of a high-fidelity VR world.

Similarly to E6, the context of the VR world was visible by increasing the brightness of the HWD to its maximum setting without the HWD being fully occluded. An example setup of this display system with its associated support cables and testing environment is shown in the following Figure:

Figure 41: E8 Display Experiment Example

In addition to these support cables, several custom hardware brackets were designed and 3D printed to allow for additional capabilities in the HWD system, as is shown in the following Figure:

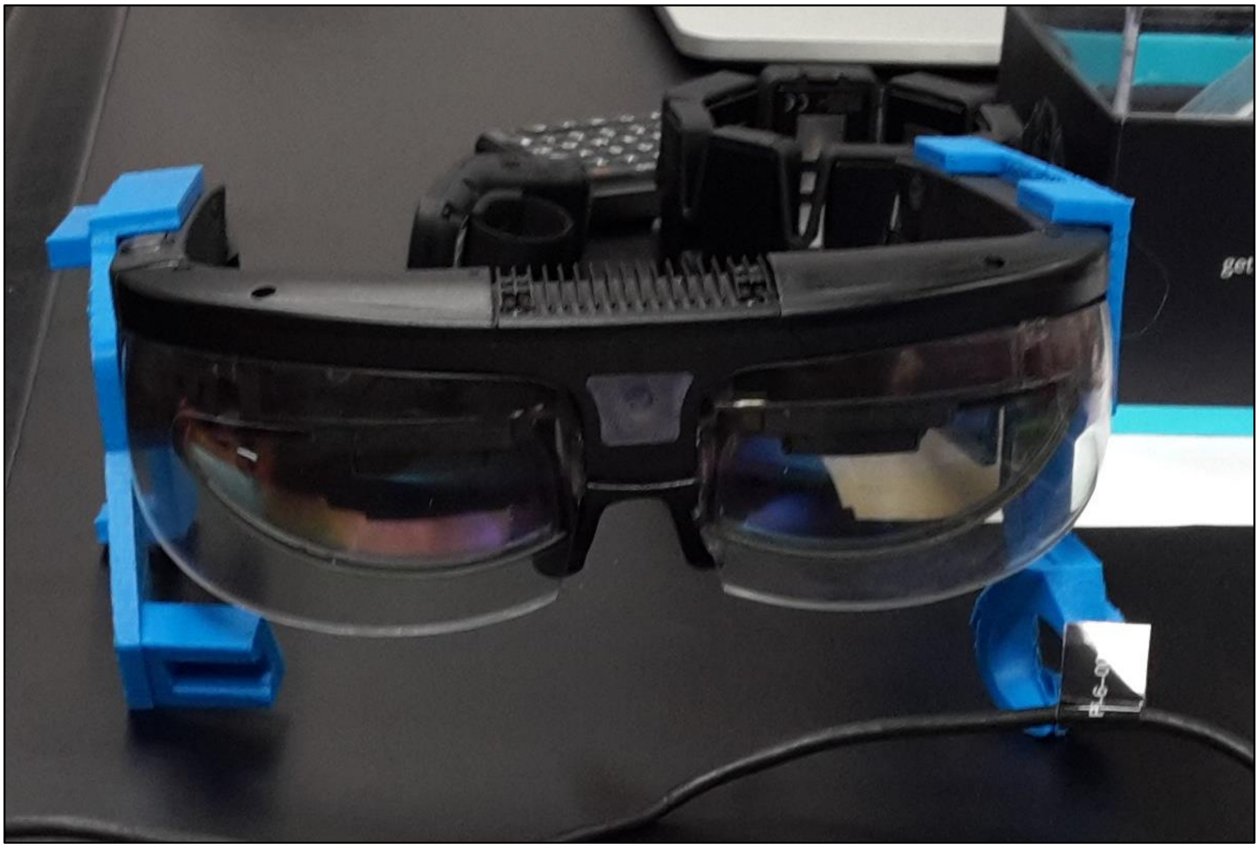


Figure 42: Creating Custom Hardware

Finally, a set of user interaction and control devices was designed and built as well, including a touch-sensitive glove, shown in the following Figure:

Figure 43: Touch Glove Device

Additional consumer-level gesture tracking, haptic feedback, keyboard, and mouse devices were selected to integrate into the user interaction testbed, as appear in the following Figure:

Figure 44: Additional Interaction Devices

Each of these devices was given to the user in order to operate the user interface system. The unseen digital voice control system was also operational during E8. The participants were questioned and results were recorded. However, many technical glitches occurred throughout the experimentation of E8, so little useful data was actually collected during E8. No single user session during E8 was successfully executed to its completion; therefore no resultant data will be shared here.

5.1.1.5.2: E9: Tethered AR HWD with Real World v1

After the hardware and software failures of E8, a simplified version of E9 was quickly iterated that removed the VR environment and allowed for a display of information elements only within the real world office workspace of E9. E9 had all the same interaction capabilities of E8 and the same semi-structured interviews were conducted with first responder and technical expert users. However, each experiment in E9 also was unsuccessful in its completion; there was always a software or hardware error that required the test to be stopped, so no useful user-centered feedback was gathered during E9 apart from the general necessity to build a more reliable user interaction system. As a result of no successfully completed participant experiment, no feedback data will be shared here.

5.1.1.5.3: E10: Large FOV AR HWD with Real World

A professional-grade large FOV AR HWD was acquired by the research team during A8 that had not been part of the original design and prototype plan of this Activity. After the hardware and software failures of E8 and E9, E10 was simplified to a display-only user interface system; no user interaction system was used to operate it. This greatly decreased the complexity of the E10 experiment and allowed the research team to properly test and refine the software required for this Experience. Several dynamic information elements were displayed (e.g., compass, MR markers, time of day, head pose, system messages) in addition to a few screen-fixed information elements meant to convey general concepts (e.g., a navigation route, radio operation). While these elements had been tested before, the HWD used in this experiment was over twice the viewable FOV of previous Experiences and allowed the research team to experiment with how information elements might be reconfigured and redesigned in this larger physical display setup. The following Figure displays a one-eye perspective on this HWD of the navigation time point within the common response scenario in the assault team member profile:

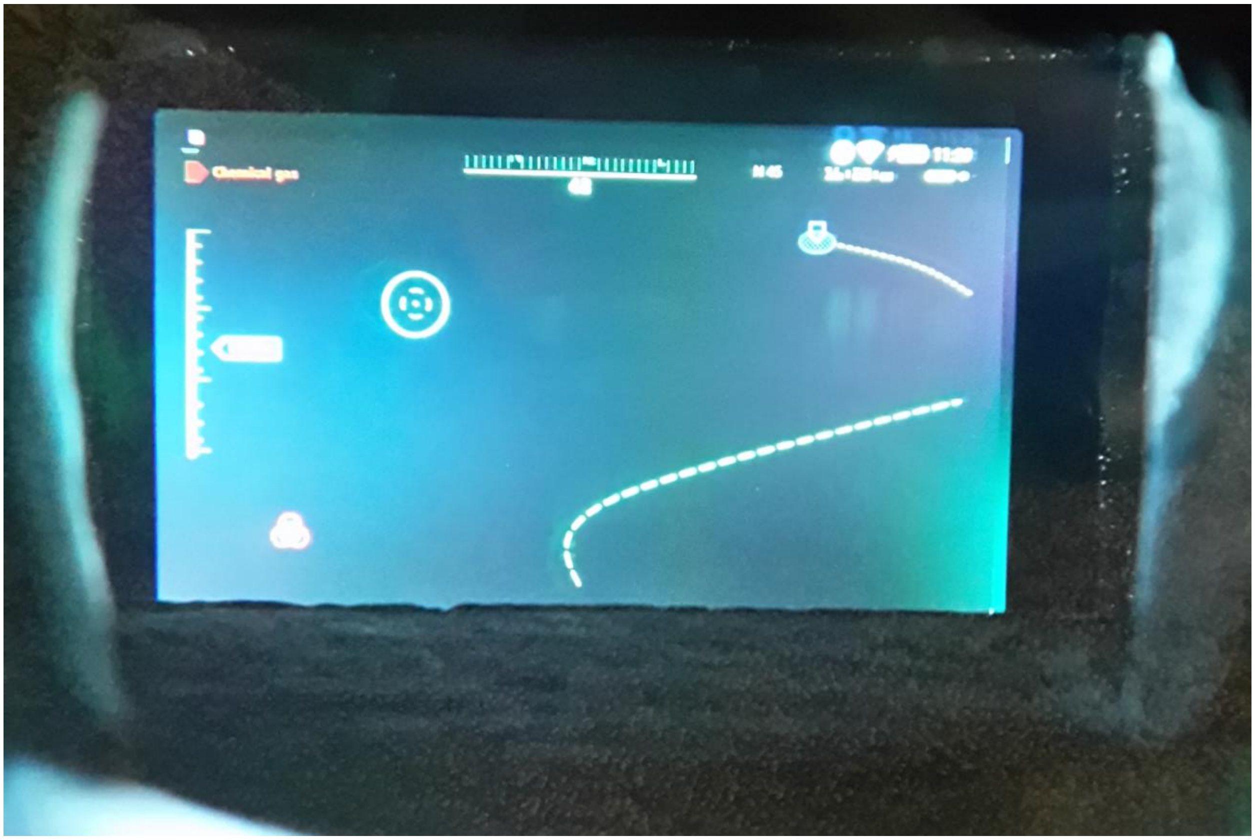


Figure 45: See-through Optical Display of E10

5.1.1.6: Discussion

Three interaction Experiences were created during A8. Many iterations were made within each of these Experiences in order to find an appropriate tradeoff in terms of what was feasible (e.g., in terms of commercially-available resources) and possible (e.g., in terms of rapid prototyping, short iteration time scales) to accomplish. While E8 included state-of-the-art, custom-built hardware and software systems, the hardware and software was utterly unreliable for experimentation with first responder participants and did little to actually progress the research effort overall. When a participant was placed in the testing scenario with all of the user interface and user interaction capabilities that had been designed into the system, the research team was unable to predict how many time points could be explored with the participant. Additionally, once the system inevitably crashed or failed to respond, the research team was unable to salvage the test without resetting the entire system. No end-user test of all five time points of the common response scenario was ever recorded.

Using the VR-based E8 display system represented a step back in user-perceived progress as well, but the AR-based version of the same experience, E9, also had to remain tethered to a single location in order to allow user-based interaction to occur or to provide any storyboard context to the experience. Without a VR environment, there was very little visual context provided to the user in terms of experiential immersion because the test had to be conducted indoors with a power supply and several computers to operate it. Users expected, at this stage of the overall research effort, that everything would be wireless and they could walk anywhere they wanted to experience A8.

With the removal of the VR world context in E9, the same software and hardware problems plagued the research team. No comprehensive test including all five time points of the common response scenario was ever recorded with a non-research team participant. The hardware and software of E9 was still too complex and unreliable to execute a reliable experiment. One particularly challenging difficulty during E8 and E9 was from a commercially-available armband device that was meant to track arm-level gestures. Researchers determined through a series of internal software testing that this device only worked on participants that had very little arm hair. If you had hairy arms, it would not function at all.

After the hardware and software systems of the previous two Experiences, the E10 system utilized a professional-grade larger FOV tethered AR headset. While the tethered HWD remained another common complaint across first responder participants, the physical tether allowed for this larger FOV to be possible in the first place. Although this detracted from a futuristic 3D MR HWD capability that was being designed for first responders, the software systems built for E10 were extremely reliable and every semi-structured interview with participants was completed successfully.

The largest problem experienced during this Activity was the custom hardware and software built for the research testbed. While highly competent developers were used to create the needed software systems – which only worked consistently during internal testing – once the systems were deployed to a user, they failed to function for myriad reasons that could never be remedied one-hundred-percent of the time. Some of the reasons for which the system would fail include the following:

- Voice-recognition subsystem:
    - Training often required to recognize user's voice
    - When trained, user was required to give several duplicate commands before the system would recognize the command
    - Commands were often not recognized after several attempts, resulting in user frustration
- Gesture/haptic subsystem:
    - Users with hairier arms were not recognized at all
    - Gesture training was required
    - Gesture motion required exact movement patterns; very low tolerance for variation
- Interaction control subsystem:
    - Bluetooth devices often powered off and failed to reconnect to software system
    - Touch glove was sensitive to becoming detached from hardware cabling
    - Complicated button controllers required user training
- HWD subsystem:
    - High-power requirements for E8-9:
        - HWD had to be plugged in at all times
        - HWD often overheated and system froze

The unreliability of these systems resulted in very low levels of interaction and user interface immersion; the antithesis of dedicating scheduled time and resources to developing a high-fidelity prototype in the first place. After these failures, it was essential for the research team to reevaluate the current trajectory of the effort and re-align the effort goals with the practical execution of UCD-based experimentation for a 3D MR HWD interaction system.

5.2: Activity 9: Action Elicitation

While A8 helped to develop several prototype systems, these prototypes did little to create an immersive user experience overall due to hardware and software inconsistencies. A9's purpose was to re-evaluate the outcomes of A8 and create user-centered design documentation and concepts that could be implemented into an interactive prototype and tested with first responders. It was also essential that the output of A9 would be reliable and provide useful UCD feedback to the research team. Finally, a series of experiments were performed with a large pool of first responders to address the RQs of O3 to a reliable degree of certainty. The following Figure communicates the methods employed during A9 and the output research artifacts that were made during it:

Figure 46: A9 Overview

5.2.1: Method

5.2.1.1: Participants

One first responder student joined the research team for A9 and served as the representative voice for this Activity's design process. A set of three first responder students were then used in a pilot study for A9. 23 first responder students were interviewed individually in a structured interview session for the full-sized study during A9. Each first responder student was a member of the Corps of Cadets at Virginia Tech. None of the students used in A9 had prior exposure to the research effort.

5.2.1.2: Instruments

5.2.1.2.1: E11: Projected Elements on a Projected Environment v2

Whiteboarding, digital sketches, graphic editing software, and video presentation instruments were used in A9. A more physically and visually immersive environment was then created in a laboratory office building, including the following characteristics of the common response scenario:

- Full-screen projected situational environment (simulated by a video game)
- Testing protocol script (to establish the storyboard and testing scenario)
- Gesture-based video recording devices
- Audio recording devices
- Immersion tools used by the participant:
    - Tactical vest
    - Ballistic helmet
    - Airsoft rifle

The following Figure is useful to communicate the physical organization of the laboratory environment, including the dual projector setup in relation to the test subject and research team computing devices:

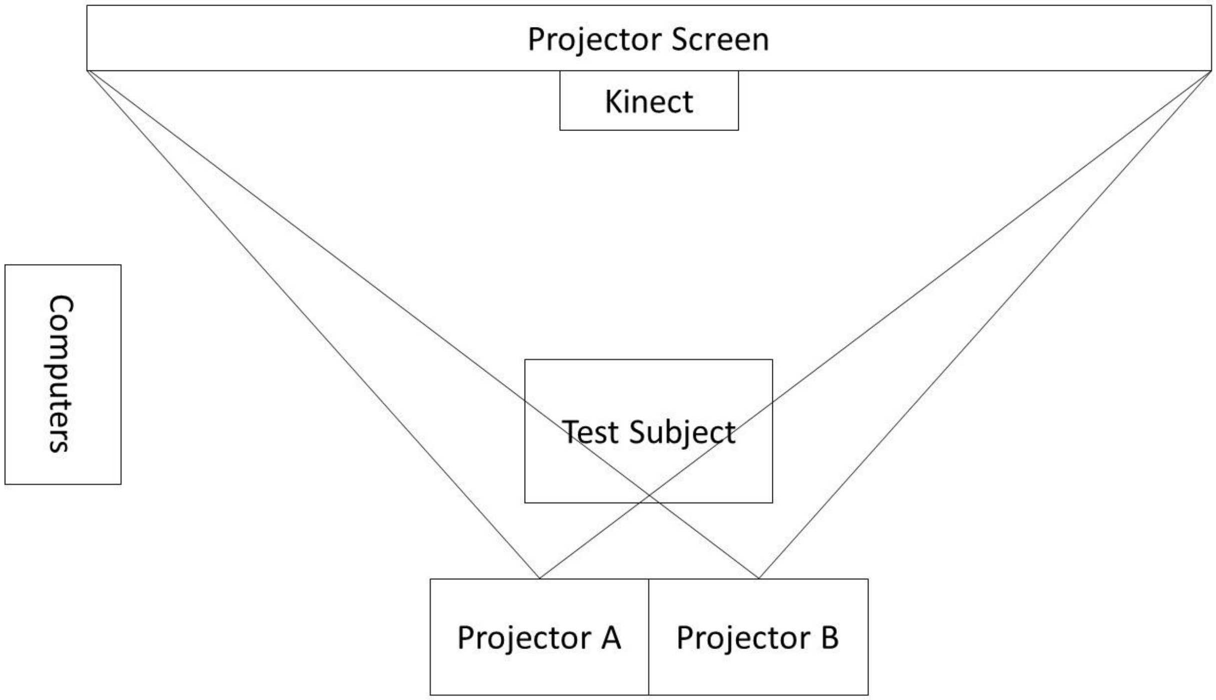


Figure 47: E11 Laboratory Environment Setup Diagram

Video, audio, and skeletal tracking recording devices were used so that a post hoc Content Analysis could be performed by the research team. The results of the test were cataloged in a spreadsheet format for later review. The display hardware utilized two traditional projectors, each with a resolution of 1080p that was shown on a 6.5’ screen at a distance of 6’ from the participant’s head.

5.2.1.3: Procedure

5.2.1.3.1: Pilot Study

The participant was first required to sign an Institutional Review Board (IRB) release for this and the remaining Activities of O3 that included student participants. After donning the immersive gear provided by the research team, the student was introduced to the common response scenario and was instructed to provide concurrent think-aloud feedback for all of their actions. These instructions were provided verbally and displayed on the wall in front of them.

A set of demographic information was gathered from each participant (e.g., name, age, technical expertise level). Once ready, three priming questions were given to the participant to introduce them to the prompts that would be used in the test. These questions included a justification statement, followed by a command, as is shown in the following example:

- You want to communicate the number “two” to others without using words.
- What action will you perform to communicate the number “two”?

The participants were then questioned accordingly at the appropriate time point of the experiment to perform the following actions:

- Select and hide an element
- Select and cycle an element
- Clear and restore their screen
- Activate and deactivate system
- Send a text message

Before each question, the research team would move the video game character to a new position in the virtual environment. Each type of question was asked at least twice during the experiment. At the end of the test, each participant was asked what they liked and disliked about E11 and what they would change in future versions. Each structured interview session was recorded via multiple devices for post hoc digital cataloguing, Content Analysis, and review by the research team.

5.2.1.3.2: Full Study

An updated version of the information elements from A8 were redesigned in A9. These elements were arranged in a HUD format for the first responder assault team member profile that had been used in previous experiments. An updated version of the common response scenario was also developed. These research artifacts are included in the Results section.

The participant was required to sign an Institutional Review Board (IRB) release for this and the remaining Activities of O3 that included student participant interviews. The subject was asked to wear a tactical vest, ballistic helmet, and an airsoft rifle that were meant to increase the realism of the testing scenario. The common response scenario was explained to the participant verbally and displayed on the wall in front of them. A real-time video game environment was also projected on the wall in front of them, which added to the realism of the scenario.

A set of demographic information was gathered from each participant (e.g., name, age, technical expertise level). Once ready, three priming questions were given to the participant to introduce them to the prompts that would be used in the test. The participants were then questioned accordingly at the appropriate time point of the experiment to perform the following actions:

- Clear and restore their screen
- Select and cycle an element
- Select and hide an element
- Select and show an element
- Send a text message
- Activate and deactivate system

Before each question, the research team would move the video game character to a new position in the virtual environment. Each type of question was asked multiple times during the experiment.

Each structured interview session was recorded via multiple devices for post hoc digital cataloguing, Content Analysis, and review by the research team.

5.2.1.4: Summary

The video and audio results of each participant were reviewed by the research team and converted into a spreadsheet data format that catalogued each first responder’s actions. A Content Analysis was performed. No time limit was enforced during any interview, no participant withdrew from E11. Most tests lasted approximately 15-20 minutes. A prototype experimental testbed was designed and tested in A9.

5.2.1.5: Results

5.2.1.5.1: Pilot Study

As explained earlier, A9 includes many new and updated research artifacts designed to reevaluate the progress that had been made in A8. The common response scenario was changed slightly to be more directly applicable to the situational environment of the experiment and to include the context shown in the following Figure:

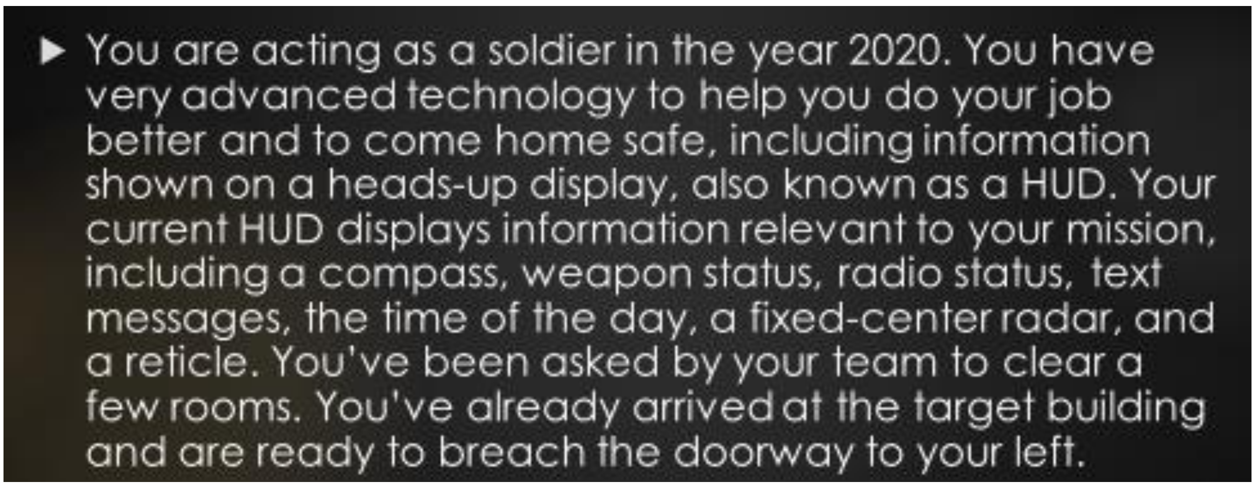


Figure 48: A9 Pilot Study Scenario Overview

A newly updated set of information elements was also designed to show a higher-resolution set of more complex data representations. These were largely borrowed from popular video game titles and modified as necessary to fit within the display constraints of E11. These elements were organized into a standardized HUD, as shown in the following Figure:

Figure 49: A9 Pilot Study Training HUD

Additional versions of each information element were also developed to allow for certain experiment prompts to function (e.g., select and change the element).

The following Figure shows a third-person perspective of the testing environment with the student first responder participant. The common response scenario description is shown on the wall in front of the participant and the multiple computer and recording devices required to execute the experiment are visible around the room.

Figure 50: A9 Pilot Study Testing Environment

Because A9 was primarily a design exercise, no further research artifacts were developed.

5.2.1.5.2: Full Study

Small design revisions were made to the information elements used in the pilot study of E11. The default HUD presentation for the full study is shown in the following Figure:

Figure 51: E11 Full Study Training HUD

Additionally, small changes to the common response scenario description were also made:

> You are a member of a SWAT team in the year 2050. You have very advanced technology to help you do your job better and to come home safe, including information shown on a heads-up display, also known as a HUD. Your current HUD displays information relevant to your mission, including a compass, radar, radio status, text messages, weapon status, the time of the day, and a reticle. You have been asked by your team to clear the compound. You have already arrived at the target building and are ready to breach the doorway.

The following Figure shows that the environmental immersion and other supporting technologies used in the pilot study of E11 were otherwise unchanged for the full study:

Figure 52: E11 Full Study Environmental Immersion From the Third-person Perspective

The first-person perspective of the participant appeared as shown in the following Figure:

Figure 53: First-person Perspective of Training HUD in Full Study of E11

E11 provided myriad data-based research artifacts. Three priming gestures were elicited by the research team in addition to 17 generic actions in each interview. Participant actions were classified into four general categories:

- Gesture
- Verbal
- Physical
- Eyes

Gesture actions generally utilized a large body motion, like waving a hand or arm in front of the participant's body. Verbal actions generally included a speech command intended for the HWD system. Physical actions generally involved pressing a hardware button of some kind. Eye actions generally were intended to activate an eye-tracking system within the HWD device. For standardization across all subjects, any action directly utilizing an eye as a means of user interaction (e.g., looking at a specific information element and intentionally blinking at it) was considered an eye action. When more than one category of user interaction was clearly used, this was noted for post hoc research analysis as a "Combo" category. The following Figure shows the frequency with which each of the four categories of user interaction were utilized during the full study of E11:

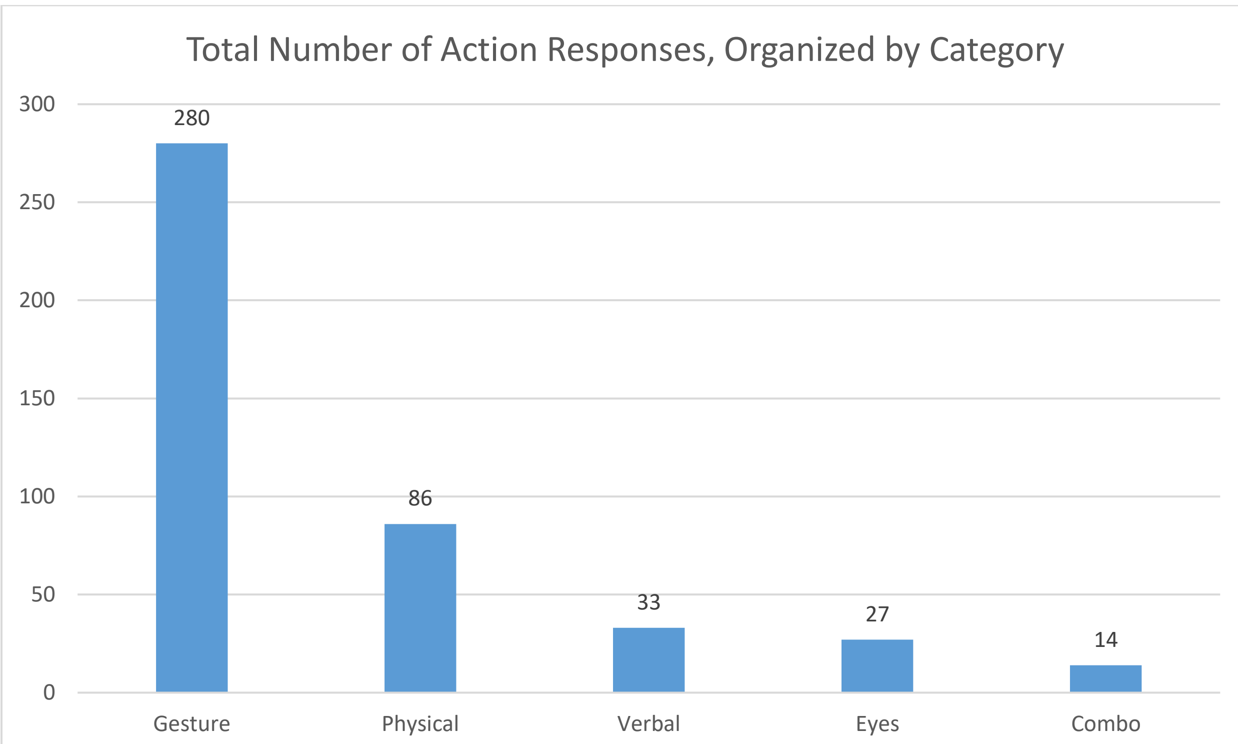


Figure 54: Total Number of Action Responses, Organized by Category

The data indicates that the vast majority of actions performed (n=280) used a gesture-type response. When graphed in a pie chart format, the following Figure shows the percentages of each category response:

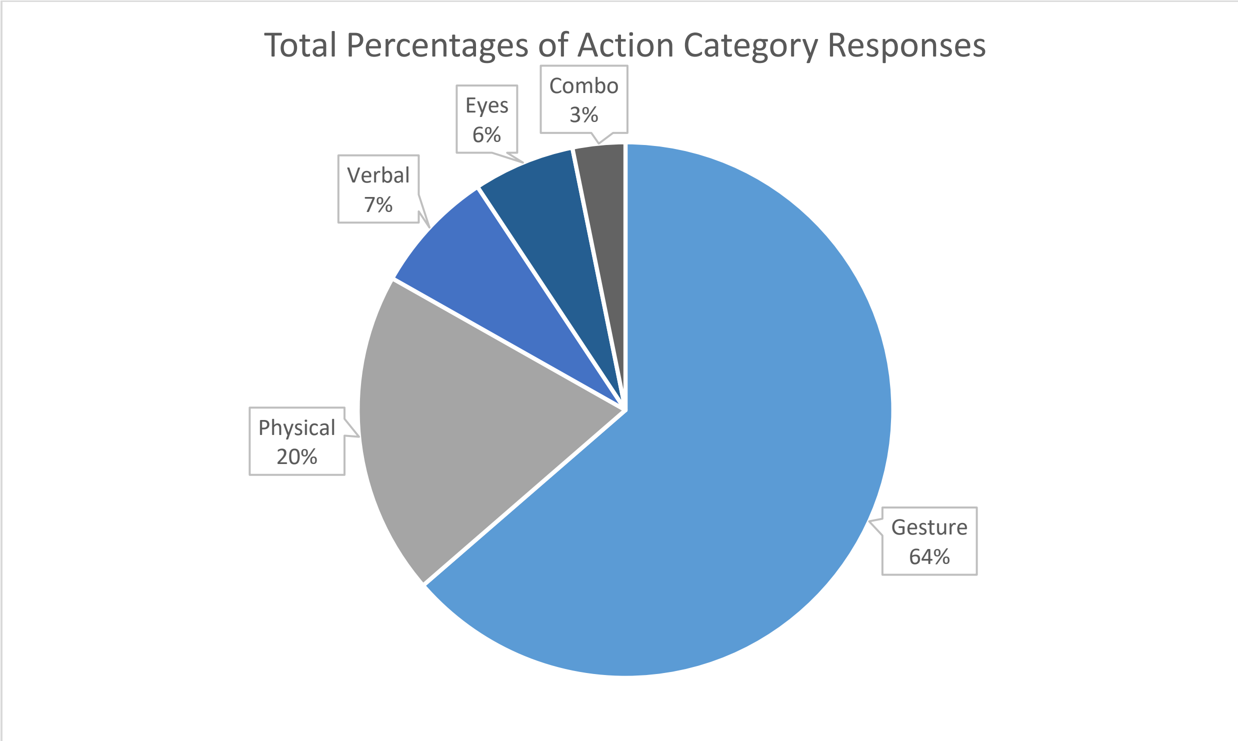


Figure 55: Total Percentages of Action Category Responses

While 64% of all action responses included a gesture, when paired with the second most popular response, Physical, the two top categories of action responses incorporate 84% of all responses. To further break down the data, the following Figure describes the number of action category responses organized by each of the 20 total questions from the experiment:

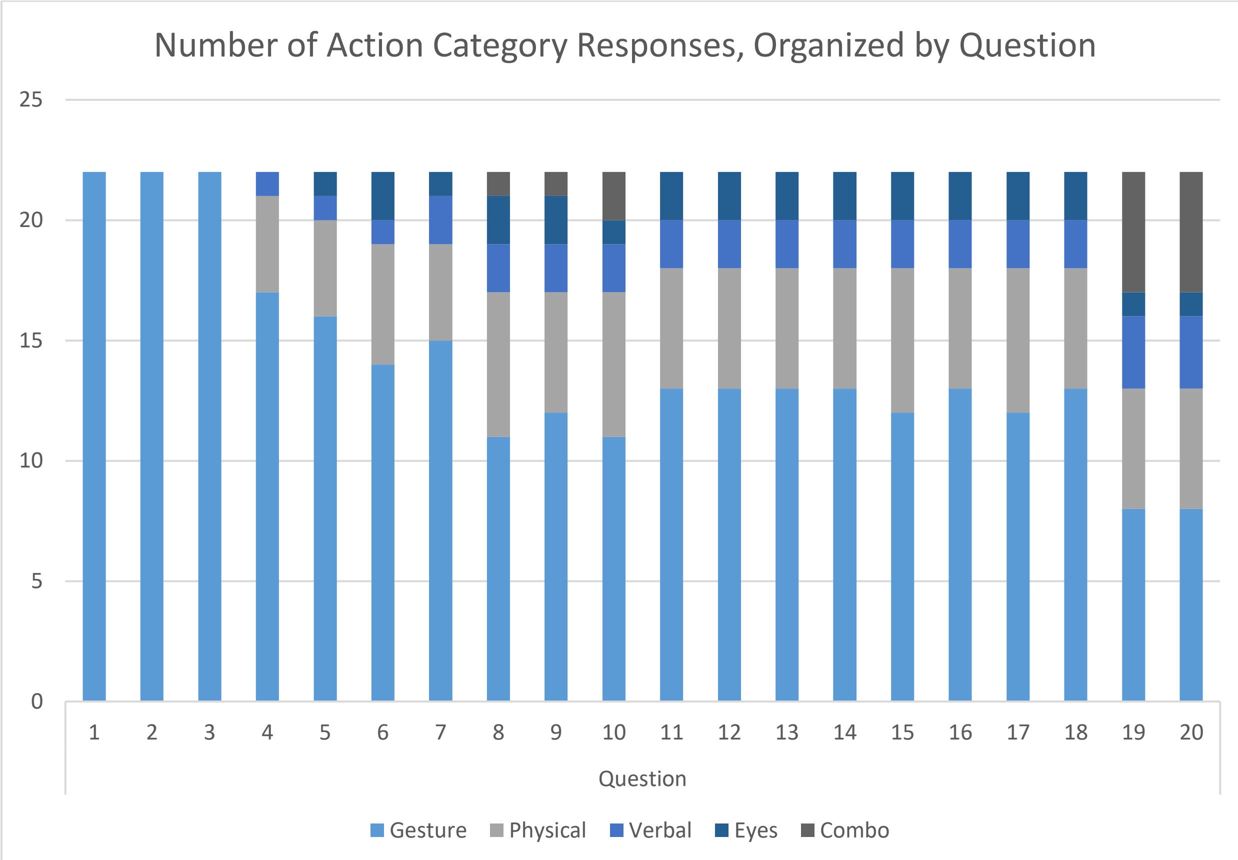


Figure 56: Number of Action Category Responses, Organized by Question

This Figure shows that in all but the last two questions, the majority of user responses fall in the gesture category. When gesture and physical categories are combined, they are always the dominant responses across E11. Additionally, the large majority of combo action responses fall in the final two questions of the experiment. The data can be additionally analyzed when the action responses are organized by each participant, as is shown in the following Figure:

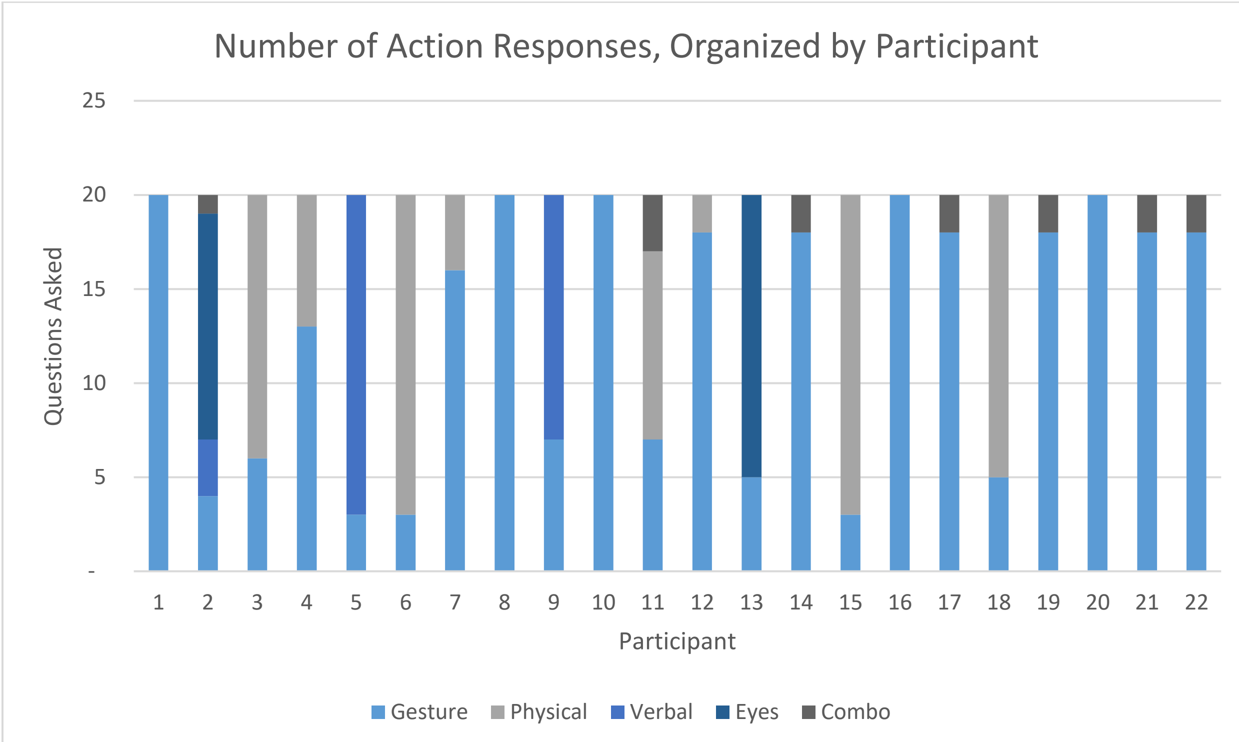


Figure 57: Number of Action Responses, Organized by Participant

While two participants relied on eye-based responses for the majority of their actions, these users represent outliers to the common trends of the remaining first responder pool.

5.2.1.6: Discussion

5.2.1.6.1: Pilot Study

Because of the failures from A8, it was necessary to design a failsafe research testbed Experience that could be quickly deployed in order to gather useful feedback on how first responders would prefer to interact with a futuristic 3D MR HWD system. While A8 attempted to physically integrate a set of user interactions, it was shown to be too unreliable to gather useful user-centered feedback. The highly-complex rapid prototype system of A8 ultimately failed. A9 was meant to take a step back, rely on previously proven and reliable PD-based storytelling methods, and design an experiment that could gather the interaction-based data that was originally designed to be gathered from A8, but could not be recorded because of highly unstable hardware and software products.

The research team considered the addition of aural simulation to E11, but decided to omit it so as to avoid any communication interference during the semi-structured interview discussions.

In order to ensure a successful first responder experimentation session, A9 included no software programming to operate and no functional user interaction hardware was utilized. It was a manually operated Experience that required two research team members to operate. Each recording device was manually started and stopped. The in-game player that was designed to represent the first responder in situ was controlled by a researcher as well. The successful execution of each experiment, therefore, was completely dependent on the competency of the research team in carefully performing each step on their own. This required much greater care on the part of the research team during each interview session, but resulted in reliable and repeatable test results.

With the data gathering process clearly defined and shown to reliably gather the necessary user-centered feedback, the research team was able to run a full-scale experiment with a large pool of first responder participants.

5.2.1.6.2: Full Study

Although 23 subjects were interviewed in total, the recording devices used to track the participant data failed during the 23rd test. Only the 22 fully recorded user data sessions were presented in the Results section and used during the analysis of A9.

A rapid iteration in design allowed for a single Experience to be explored during A9 that resulted in quantitative and qualitative data points that directly addressed the RQs of O3. While A8 was intended to address these same questions, the reliability of that rapid prototype failed to yield qualitative results to progress the overall research effort.

The largest complaint from participants was due to the lack of user-based control during this experiment – a sentiment that should sound all-too familiar by now to the reader. While every participant in each rapid prototype experiment continued to express this desire throughout the entire research effort, it was important to step back from this continued request during A9 of the overall research in order to not repeat the mistakes of A8.

From an external observer's perspective, it is also important to address the change in first responder participant pools during A9. While previous Activities utilized highly-trained SME participants that met the criteria of the ideal first responder, as described in O1, it was necessary to move to a different participant pool for many reasons:

- The research team was no longer co-located full-time with the ideal first responder participant pool used in previous Activities
- The use of PD-based storytelling methods had previously proven highly-effective in mentally situating first responder participants into the common response scenario, so it was postulated that more generalized first responders would appropriately fill the need of future experimentation session
- Corps of Cadets members have training similar to other first responder groups (e.g., police, paramedics) that perform the common response scenario regularly; they simply represent a more novice subset of these same groups
- Corps of Cadets members might have greater levels of familiarity with future technologies (e.g., play more video games, study more future trends, more prone to ideation)
- As defined initially in this research paper, Corps of Cadets members clearly fall under the same umbrella definition of "first responders"
- Corps of Cadets members were available and eager to participate in experiments
- Corps of Cadets members represent a greater representation of the general population of first responders instead of the highly-specialized SMEs that were relied-upon in early Activities
- One of the conclusions of the expert SMEs in O1 was that the common response scenario was an action familiar enough across the security-based first responder domain, that it would apply to a more generalized domain, like that of the students used in A9.

It is important to also note that the number of measured responses (i.e., 440 actions) from the participant pool (i.e., 22 users) represent a very low level of statistical variation. Only two participants utilized an eye response at all; three used a verbal action. Half of the participants utilized physical actions. It is expected that in future experiments of this type, that more specialized

SMEs would follow similar trends with the vast majority of them utilizing gesture and physical action responses to these same question prompts.

It is important to note a slight variation to the Gesture Elicitation method described in Chapter 2 of this paper. While the Gesture Elicitation technique for gathering useful UCD feedback was followed generally, this Activity was designed to explore multi-modal methods of communication that might be most natural and intuitive to the first responder participants interviewed throughout this Experience. While Gesture Elicitation focuses on gestures alone, any conceivable type of action response from the participant pool was considered a valuable response to the overall research effort. While Gesture Elicitation, as practiced by its current experts, is focused on gesture-based development, the more general action-based prompts utilized in this Activity provided a more broad response set than would have been achieved by a strictly gesture-based interview prompt.

Confident that an understanding of the actions participants expect to perform in a futuristic 3D MR HWD interaction system was established, A10 addresses the ongoing desire to have more control over the interaction testbed and a renewed attempt to more carefully implement user interaction control devices into the experimental testbed.

5.3: Activity 10: User Controls

While A9 helped to create an Experience that could reliably be tested with first responder participants during a fast, iterative experiment, A10 was used for a final user interaction prototype system that could be controlled by the participant themselves. This included myriad rapid design and prototype iterations over 12 months of time to develop a reliable set of Experiences. The following Figure shows the methods used and research artifacts developed during A10:

Figure 58: A10 Overview

5.3.1: Method

5.3.1.1: Participants

The same first responder student participant from A9 served as the voice for A10 in partnership with the research team. 10 first responders were interviewed in a semi-structured manner to gather feedback on the final versions of each interaction Experience.

5.3.1.2: Instruments

Whiteboarding, digital sketches, and presentation instruments were used for design iterations. Designs were created for a newly released untethered see-through AR-based HWD. Various software and development suites were used to construct each Experience. E12 and E13 formal experimentation was performed in a similar laboratory environment to E11, including the tactical gear worn by the participant and a projected virtual immersion environment. More details on each Experience are described in the following sections:

5.3.1.2.1: E12: On, Off, Cycle

A newly released untethered AR HWD was acquired by the research team. This HWD used existing development tools that made the creation of E12 relatively easy to deploy compared to previous Experiences. The research team targeted a limited set of user interactions to execute in E12. The user was provided the training HUD from E11 and could use the gaze tracking of the HWD to turn any information element on, off, or to cycle through a set of elements that would change the specific visualization of that piece of data. A single Bluetooth-based game controller was used to perform these three interactions with a limited set of controller buttons. Three versions of E12 were ported from the same basic code base, but resulted in slightly different experimentation characteristics:

- E12a: Untethered AR HWD w/projected environment v2
- E12b: Mobile phone w/low-fidelity VR world v2
- E12c: Mobile phone w/real world v2

The same video game virtual environment of E11 was used again in E12 to provide additional context to the participant in the laboratory-based experiement. The E12a display hardware utilized a traditional projector with a resolution of 1080p that was shown on a 6.5’ screen at a distance of 6’ from the participant’s head. The HWD hardware utilized a Windows-based device with a binocular optical see-through display with a resolution of 720p with a total FOV of 35-degrees diagonal. The E12b display hardware utilized an Android-based smartphone with a resolution of 445p in a 5” size that was shown in a VR headset with a 96-degree diagonal FOV. The E12c display hardware utilized an Android-based smartphone with a resolution of 445p in a 5” size that was shown at a variable distance of less than 2’, depending on where the participant held the device.

5.3.1.2.2: E13: Menu Comparison

The same novel AR HWD of E12 was used here, along with the same game controller interaction device. The user was allowed to interact with E13 by using two different designs of menu systems: radial and linear organization. A pre-recorded video game virtual environment was played for the participant in place of the researcher-controlled environment of E12. A screen shot of this video file is shown in the following Figure:

Figure 59: E13 Video File Screen Shot

The same three versions of E12 were replicated for E13:

- E13a: Untethered AR HWD w/projected environment v3
- E13b: Mobile phone w/low-fidelity VR world v3
- E13c: Mobile phone w/real world v3

Video and audio recordings were made for later analysis and categorization of user feedback. The E13a display hardware utilized a traditional projector with a resolution of 1080p that was shown on a 6.5’ screen at a distance of 6’ from the participant’s head. The HWD hardware utilized a Windows-based device with a binocular optical see-through display with a resolution of 720p with a total FOV of 35-degrees diagonal. The E13b display hardware utilized an Android-based smartphone with a resolution of 445p in a 5” size that was shown in a VR headset with a 96-degree diagonal FOV. The E13c display hardware utilized an Android-based smartphone with a resolution of 445p in a 5” size that was shown at a variable distance of less than 2’, depending on where the participant held the device.

5.3.1.3: Procedure

The student first responder was asked to create several updated versions of the information elements of the updated common response scenario. The feedback from A9 was used to create more legible data representations that could be displayed on the untethered AR HWD. Six computer science undergraduate students joined the research team during A10 to build and test E12 and E13. The Experiences that were developed focused on a single user interaction request: to customize the training HUD. The following Figures represent the results of one of the design sessions from E13:

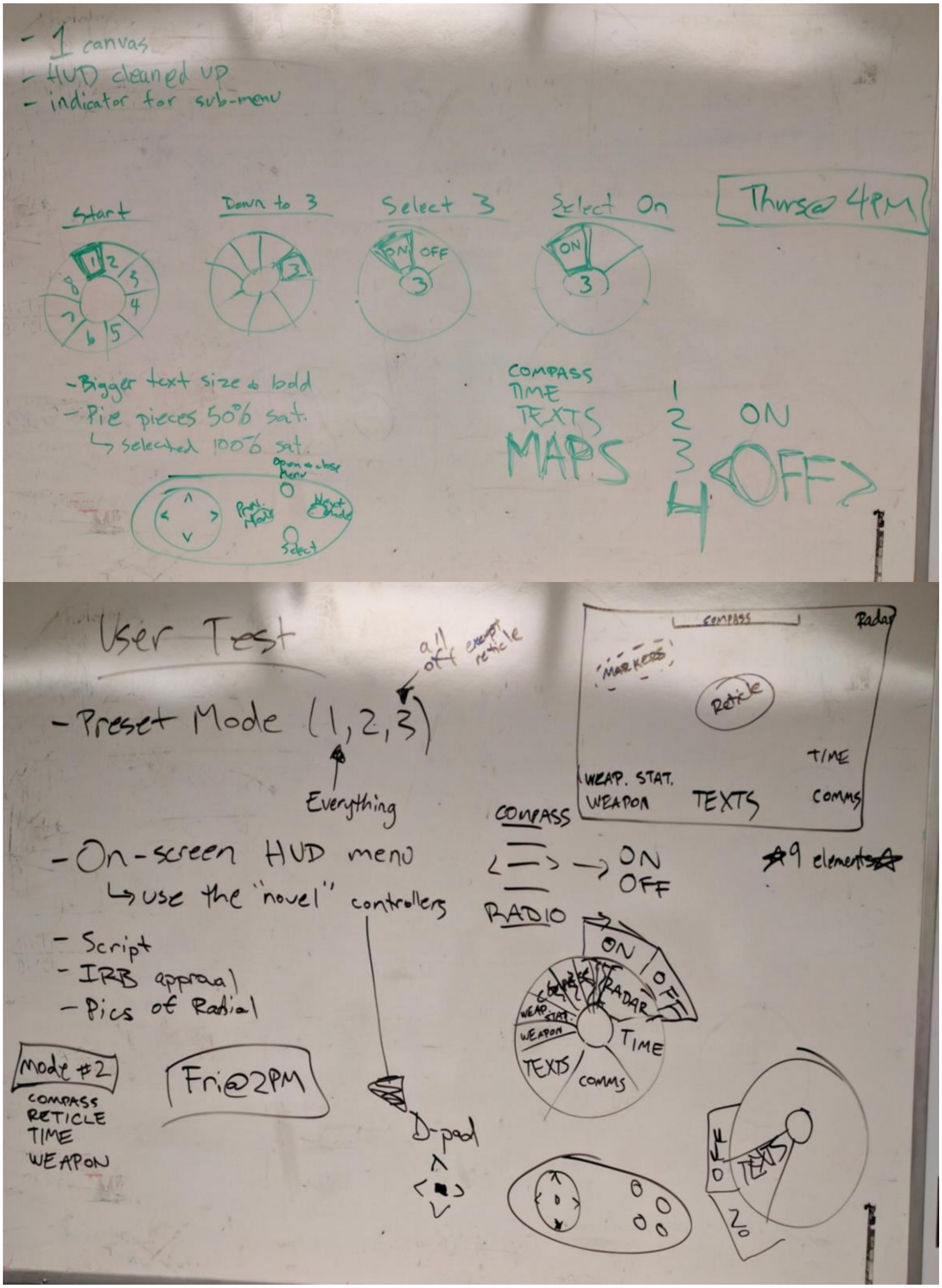


Figure 60: E13 Design Session

PD-based storytelling provided the context of the common response scenario and the relevant user role from A9. Each participant was asked to provide feedback on each method of interaction with the futuristic 3D MR HWD system. Each semi-structured interview began with an orientation of the available controls (e.g., select, cycle). The participant was then allowed to customize their HUD as much or as little as they preferred. The following questions were then asked of each participant in regards to E12:

- Do you have any overall thoughts about this interaction experience?
- What about this experience made it easy or hard to use?
- Which of the three modes did you prefer to use?
- What would you change in future versions of this experience?
- Any final thoughts about this experience?

The semi-structured interview questions for E13 were as follows:

- For each menu design:
    - Do you have any overall thoughts about this menu design?
    - What about this menu design made it easy to navigate through the items in the menu?
    - What about this design made it difficult to navigate through the menu?
    - What would you change about this menu design, if you could?
    - Which of the three modes did you prefer with this menu design?
    - Any final thoughts on this menu design?
- After both menus were presented:
    - Which of the two menu designs did you prefer?
    - What about that menu made it easier to navigate than the other style of menu?
    - Was there anything that made both menu designs easy or difficult to use?
    - Did you prefer a certain mode with each menu type?
    - Any final thoughts on either of the menus, or anything about the experiment?

The results of each participant were converted into a spreadsheet data format in real-time during each participant interview.

5.3.1.4: Summary

20 total first responder interviews were performed. A post hoc data analysis was performed within the research team to determine trends among the participant pool. Future areas for research and development were discussed.

5.3.1.5: Results

One consistent complaint from first responders during this research effort had been the lack of control over the testbed experience. In order to address this concern, the research team designed E12 and E13. These two Experiences were determined by the research team to be relatively easily implemented and somewhat more reliable due to the release of a new, fully integrated, untethered, programmable AR HWD system. The training HUD designed and experimented with during A9 was revised to match the constraints of this novel HWD during a rapid design iteration that addressed the following constraints:

- Element colors that were clearly displayed
- Element sizes that were easily identified
- Element resolutions that were clearly displayed
- Interaction motifs that were easily executed
- Environments that could emphasize the common response scenario
- Experiments that would be reliable
- Experiments that answered the RQs of O3

The following Figure displays the first person perspective of E12a in operation:

Figure 61: E12a First Person Perspective

The next Figure displays the first person perspective of E12b in operation:

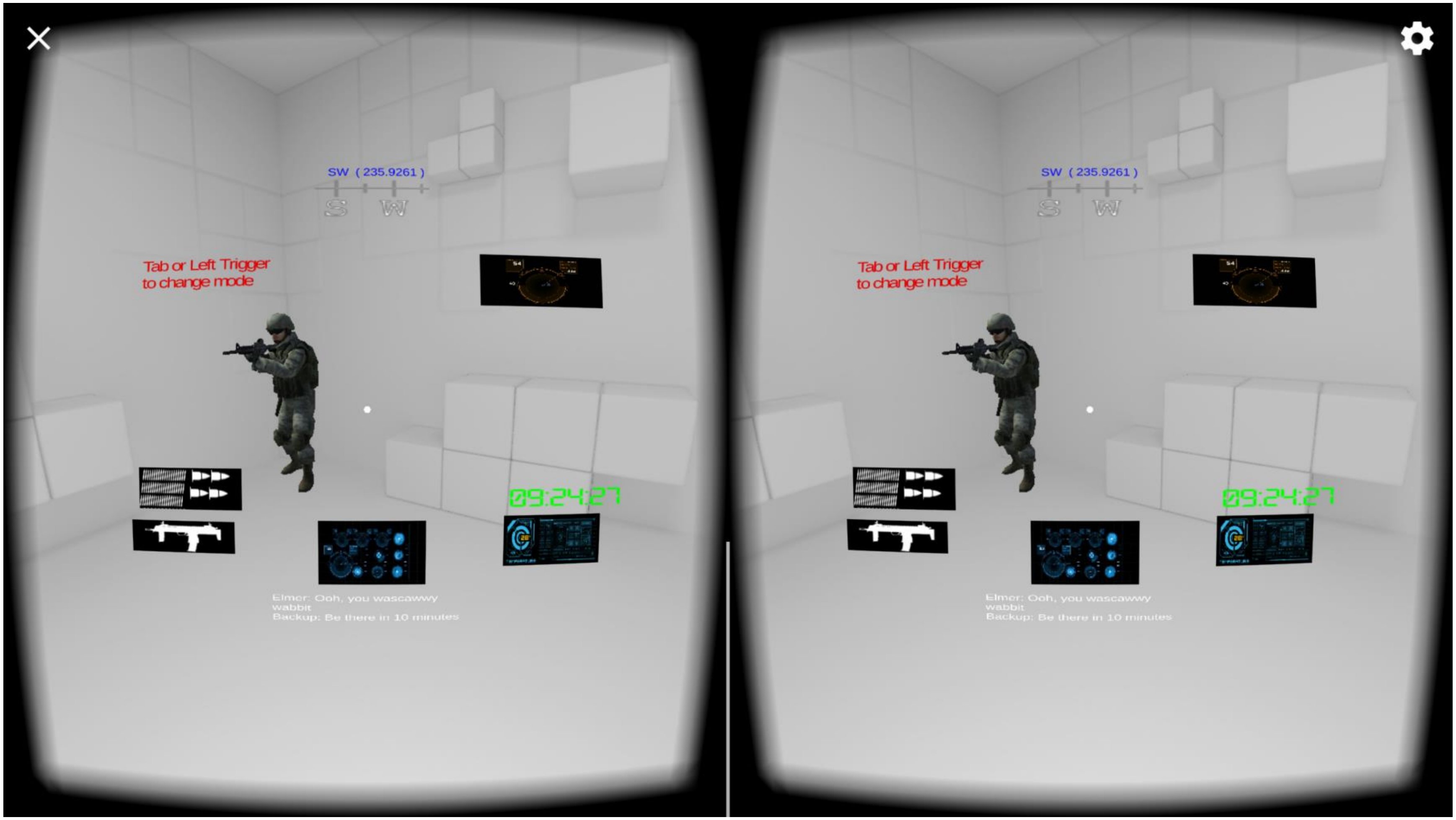


Figure 62: E12b First Person Perspective

The next three Figures show the first person perspective of E13a in the training HUD configuration before the menu system is activated, when the linear menu is displayed, and when the radial menu is displayed:

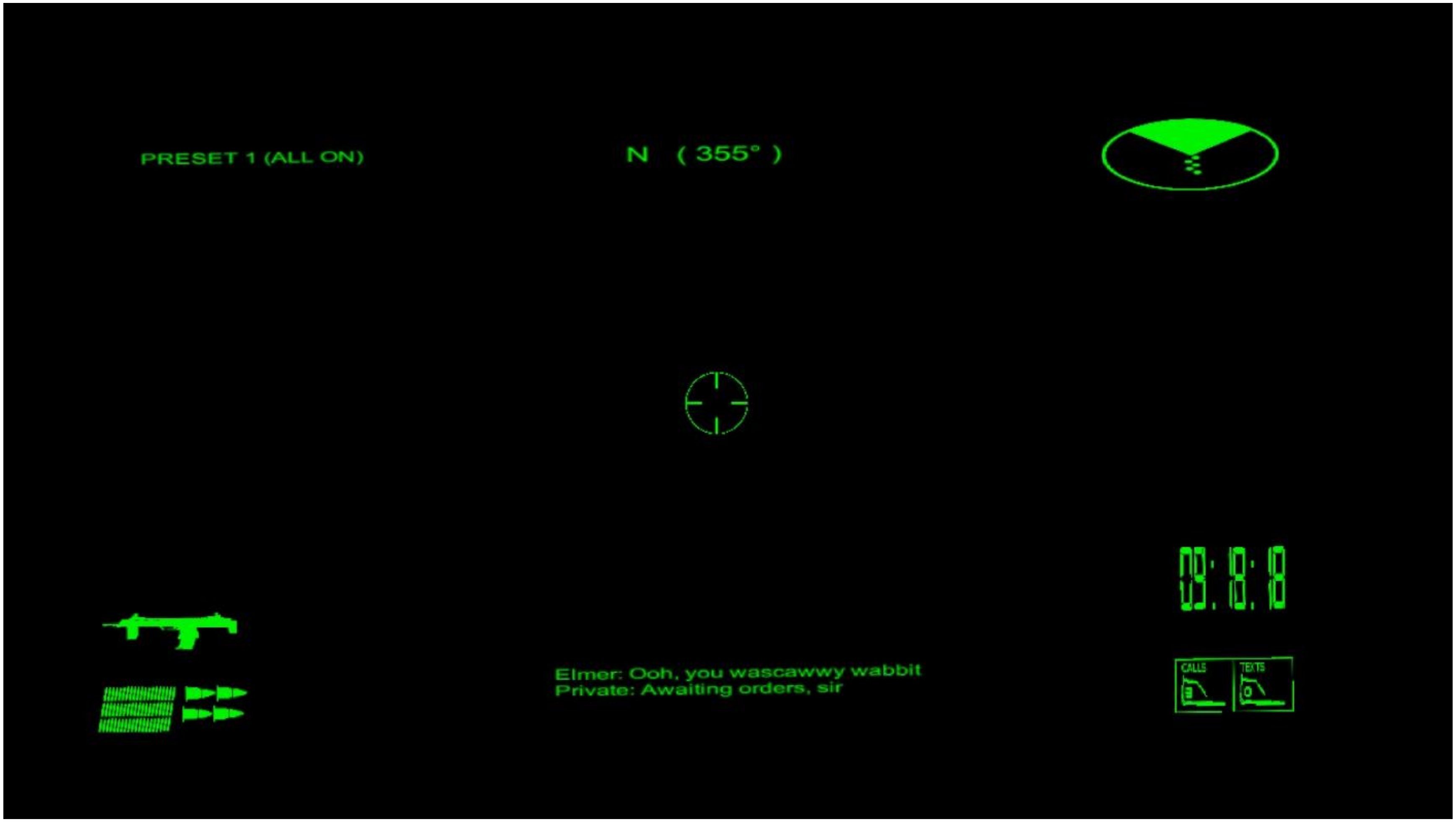


Figure 63: E13a Training HUD

Figure 64: E13a Linear Menu

Figure 65: E13a Radial Menu

The next Figure shows the first-person perspective of E13c:

Figure 66: E13c Training HUD

The resultant data from E12 was consistent in its feedback from users. Because the responses were largely qualitative, the following key take aways apply:

- Easy to customize
- I'd like to change some things more quickly
- Additional features desired
- Elements visually recognizable
- Elements visually clear
- Elements visually accurate

E13 included 10 total participants. 8 of them preferred the radial menu design, although both versions were easy to navigate.. All of the users requested that additional features be added in future versions.

5.3.1.6: Discussion

Although there were three separate versions of the E12 and E13 interaction rapid prototypes that were developed, the design methods employed during previous Experiences dictated that E12a and E13a would prove most useful toward answering the RQs of O3. The targeted interaction design and execution in A10 are largely believed to be the reasons for successful UCD experimental results, which present a stark contrast to the interaction design and execution difficulties experienced in A8. Although a large investment in resources was required to create E12 and E13, it did mark a logical next step in the overall progression of the rapid design and prototype of a 3D MR HWD interface and interaction system for first responders.

Most participants had relatively positive feedback during A10, contrary to many previous Activities. All participants agreed that the design of the interaction was easy to use and made the customization of the training HUD easy to accomplish. There were many requests for future features in future Experiences, such as a "quick setting" function and more complete control over the entire HWD system from the given interaction device. Additional interaction controllers were also requested as a logical next step in the development process with user's expressing the intention that they would use smaller button controllers mounted on their person in various locations so they could be more fully immersed in the common response scenario environment.

As might be expected at this point in the development process, it is also important to note that A10 participants had no suggestions for how to improve visual information elements; they were all cited as being recognizable, clear, and accurate in the common response scenario.

Moving forward, it would be expected that the same level of reliability would exist in future Experiences and Activities and that the basic framework established in A10 would be incrementally improved in the ways described herein. No major issues were encountered during A10, which represented a welcome outcome to the research team, especially after the difficulties encountered during A8.

With each Objective described in detail, an overarching discussion is in order for the entire research effort that compares the lessons learned and methods exercised throughout this paper. A multi-dimensional Experience characteristic analysis will also be described in the following chapter in order to draw logical conclusions that apply across all of the Objectives of this research effort.

6: Final Discussion and Conclusions

6.1: Characteristics of Experiences

While many research studies focus on a small sample of experiments, with 13 unique Experiences, this research effort covers a much more rich set of experimentation data than is typically published. Longitudinally speaking, these Experiences incorporate a 38-month duration effort that includes a depth and breadth of data that are unique to this domain, as has been explained earlier (Hornbæk, 2006; Law & van Schaik, 2010). When viewed in a longitudinal manner, the results from one experiment serve as input to the next, thereby creating an extensive knowledge base of information for both end-user participants and the research team to reference at every stage of development. It is expected that better overall user prototype Experiences can be created with this longitudinal approach. Additionally, participants that travel along the development journey with the research team, who also have some level of longitudinal understanding, are expected to be more useful end-users when utilized within Experiences that exhibit varying levels of similar characteristics. What is not understood in the current literature on longitudinal UCD-based research is how user-based priming might effect UCD-based participant feedback. Additionally, it is not understood how the varying levels of experimental characteristics might influence the participant feedback that is gathered. For example, within this domain of rapid design and prototype 3D MR HWD systems, is it better to give some degree of priming to participants (in the form of pre-experimentation education of some type or first-hand experience as a previously utilized participant?) or to only use a new set of participants for every experiment?

In order to provide a series of educated conclusions with regards to the Experiences described throughout this paper, specific criteria need to be established through which meaningful cross-experiment comparisons can be made. Although 3D MR HWD criteria are not well established, several authors have begun to discuss this topic and develop general metrics that apply to this domain. The following Table provides a foundation for a rich discussion of the differences between Experiences across the entire research effort. While previous chapters have addresses low-level differences within a single Activity, the remainder of this chapter focuses on specific characteristics of each Experience and how they compare to each other. While many comparisons and analyses can be performed due to the breadth and depth of this longitudinal 38-month research effort, a selected portion of the most salient points, as determined by the research team, will be discussed. The following categories have been chosen to compare here:

- Cost
- Interface
- Interaction
- Immersion
- Display Hardware

The *cost* characteristic represents an average of the development cost required to create each specific Experience from the beginning. Due to the in-depth notes and records taken by the research team, these averages represent the actual number of days that were required to build the Experience, multiplied by the actual number of people working on the Experience, multiplied by an average annualized salary of $60,000 for an entry-level software programmer or user experience researcher, which represented the majority of the research team ("Salary," 2017a, "Salary," 2017b). A mathematical example of the formula used for the average Experience cost is as follows:

- 2 days * 2 researchers * $240 per day = $960 to develop E1

Actual real-world costs for each Experience vary dramatically in all research applications according to the scope, schedule, and resources dedicated to the development of an Experience. The skill level of each team member is also an important factor to consider in future research efforts. However, the average costs for each Experience within this specific effort are consistently and accurately counted. It is important to iterate again that the given costs represents the entire development of the Experience from its inception point. Although many Experiences were incrementally built on previous work, in order to accurately compare each Experience, it must stand on its own and not simply represent the delta of cost between the original Experience and the later version with additional development work.

The *interface* category is defined to include the following characteristics:

- User Input (i.e., the ability of the user to perform an action that influences the interface)
- System Output (e.g., the ability of the system to perform an action that influences the interface)
- Software (i.e., the ability of a software program to control the connection between the user input and the system output)
- Functional (i.e. the interface of study responds to an action) (Bowman, Kruijff, LaViola Jr, & Poupyrev, 2001; Catani & Biers, 1998; "Doug's 1968 Demo - Doug Engelbart Institute," n.d.; Rudd, Stern, & Isensee, 1996; Walker, Takayama, & Landay, 2002)

While each Experience throughout this research effort have been described with many of the details related to the above four characteristics of the interface category, this chapter will treat each category as a binary option: either the Experience included the characteristic in its design or it did not.

The *interaction* category includes three of the characteristics described in the *interface* category:

- User Input
- System Output
- Software (Bowman & Hodges, 1999; Bowman, Johnson, & Hodges, 2001; "Doug's 1968 Demo - Doug Engelbart Institute," n.d.; R. P. McMahan, Bowman, Zielinski, & Brady, 2012; Tucker, 2004)

Because the *interaction* category represents a subset of *interface*, the same definitions given previously apply equally across the categories. These are represented by binary "yes" or "no" values.

The *immersion* category includes four unique characteristics, which are also represented as binary values:

- Tactical (i.e., "a flawless user interface, one that responds rapidly, intuitively, and above all reliably" (E. Adams, 2004)

- Spatial (i.e., looks and feels right)
- Strategic (i.e., "observing, calculating, deducing" (E. Adams, 2004))
- Narrative (i.e., good storytelling) (E. Adams, 2004; Bjork & Holopainen, 2004; Bowman & McMahan, 2007; A. McMahan, 2003; R. P. McMahan, Gorton, Gresock, McConnell, & Bowman, 2006; Wu, Lee, Chang, & Liang, 2013)

The final category, *display hardware*, represents a description of the physical devices that were used in each Experience in order to communicate the graphical interface or HUD. They include the following values and are used as a further descriptor of each Experience, rather than a value judgement of a specific targeted characteristic:

- Projector
- Mobile Phone
- VR Headset
- AR Headset

With these categories and definitions in mind, the following Table displays each Experience from the research effort, with its corresponding characteristics:

Table 11: All Experience Characteristics, sorted Chronologically

					Interface								
					Interaction				Immersion				
	Name	Description	Abbreviated Description	Costs	User Input	System Output	Software	Functional	Tactical	Spatial	Strategic	Narrative	Display Hardware
Experience	E1	Projected elements w/real world	proj/RW	$ 960					X			X	Projector
	E2	Projected elements on a projected environment v1	proj/proj/v1	$ 960		X		X	X	X		X	Projector
	E3	Mobile phone w/real world v1	ph/RW/v1	$ 1,440		X	X	X	X	X		X	Mobile Phone
	E4	Mobile phone w/low-fidelity VR world	ph/low/VR	$ 2,400	X	X	X	X	X			X	Mobile Phone
	E5	VR HWD w/high-fidelity VR world	VR/high/VR	$ 4,800	X	X	X	X	X	X		X	VR Headset
	E6	Tethered AR HWD w/low-fidelity VR world	AR/low/VR	$ 2,400	X	X	X	X	X			X	AR Headset
	E7	Tethered AR HWD w/projected environment v1	AR/proj/v1	$ 1,440		X		X	X	X		X	AR Headset
	E8	Tethered AR HWD w/high-fidelity VR world	AR/high/VR	$36,000	X	X	X	X		X		X	AR Headset
	E9	Tethered AR HWD w/real world v1	AR/RW/v1	$28,800	X	X	X	X		X		X	AR Headset
	E10	Tethered AR HWD w/real world v2	AR/RW/v2	$ 1,920	X	X	X	X	X	X		X	AR Headset
	E11	Projected elements on a projected environment v2	proj/proj/v2	$ 2,400	X	X		X	X	X		X	Projector
	E12a	Untethered AR HWD w/projected environment v2	AR/proj/v2	$27,360	X	X	X	X	X	X		X	AR Headset
	E12b	Mobile phone w/low-fidelity VR world v2	ph/low/VR/v2	$30,240	X	X	X	X	X	X		X	Mobile Phone
	E12c	Mobile phone w/real world v2	ph/RW/v2	$30,240	X	X	X	X	X	X		X	Mobile Phone
	E13a	Untethered AR HWD w/projected environment v3	AR/proj/v3	$27,360	X	X	X	X	X	X		X	AR Headset
	E13b	Mobile phone w/low-fidelity VR world v3	ph/low/VR/v3	$30,240	X	X	X	X	X	X		X	Mobile Phone
	E13c	Mobile phone w/real world v3	ph/RW/v3	$30,240	X	X	X	X	X	X		X	Mobile Phone

Variation in the characteristics of the Experiences exists in all of the columns apart from strategic and narrative forms of immersion. According to the given definition of strategic, the higher-level problem-solving-type actions that exist in other experiments in the MR domain (e.g., solving puzzles, answering trivia questions, outmaneuvering your opponent), were not present in the sets of Experiences described here, resulting in a "false" value for each experiment in that category. To the opposite effect, every Experience has a "true" value for the narrative characteristic because the PD-based storytelling method was constantly utilized throughout this research effort. Because the strategic and narrative characteristics do not vary across Experiences, they will be omitted from the future Tables in order to visually simplify the presented dataset. Therefore, the following abbreviated Table is presented for the reference of the reader:

Table 12: Abbreviated Experience Characteristics, sorted Chronologically

					Interface						
					Interaction				Immersion		
	Name	Description	Abbreviated Description	Costs	User Input	System Output	Software	Functional	Tactical	Spatial	Display Hardware
Experience	E1	Projected elements w/real world	proj/RW	$ 960					X		Projector
	E2	Projected elements on a projected environment v1	proj/proj/v1	$ 960		X		X	X	X	Projector
	E3	Mobile phone w/real world	ph/RW/v1	$ 1,440		X	X	X	X	X	Mobile Phone
	E4	Mobile phone w/low-fidelity VR world	ph/low/VR	$ 2,400	X	X	X	X	X		Mobile Phone
	E5	VR HWD w/high-fidelity VR world	VR/high/VR	$ 4,800	X	X	X	X	X	X	VR Headset
	E6	Tethered AR HWD w/low-fidelity VR world	AR/low/VR	$ 2,400	X	X	X	X	X		AR Headset
	E7	Tethered AR HWD w/projected environment v1	AR/proj/v1	$ 1,440		X		X	X	X	AR Headset
	E8	Tethered AR HWD w/high-fidelity VR world	AR/high/VR	$36,000	X	X	X	X		X	AR Headset
	E9	Tethered AR HWD w/real world v1	AR/RW/v1	$28,800	X	X	X	X		X	AR Headset
	E10	Large FOV AR HWD w/real world	AR/RW/v2	$ 1,920	X	X	X	X	X	X	AR Headset
	E11	Projected elements on a projected environment v2	proj/proj/v2	$ 2,400	X	X		X	X	X	Projector
	E12a	Untethered AR HWD w/projected environment v2	AR/proj/v2	$27,360	X	X	X	X	X	X	AR Headset
	E12b	Mobile phone w/low-fidelity VR world v2	ph/low/VR/v2	$30,240	X	X	X	X	X	X	Mobile Phone
	E12c	Mobile phone w/real world v2	ph/RW/v2	$30,240	X	X	X	X	X	X	Mobile Phone
	E13a	Untethered AR HWD w/projected environment v3	AR/proj/v3	$27,360	X	X	X	X	X	X	AR Headset
	E13b	Mobile phone w/low-fidelity VR world v3	ph/low/VR/v3	$30,240	X	X	X	X	X	X	Mobile Phone
	E13c	Mobile phone w/real world v3	ph/RW/v3	$30,240	X	X	X	X	X	X	Mobile Phone

To further simplify the discussion in the remainder of this Chapter, it is proposed that similar Experiences be grouped together and analyzed through a holistic research lens. To this end, the next Section of this paper will describe how each Experience will be described at a higher-level perspective. This higher-level academic perspective should also prove more effective to readers attempting to compare their own experiment testbeds to those that have been shared in this effort.

6.2: Reality-Virtuality Continuum (RVC)

While the previous sections of this chapter have largely referenced the specific Experience numbers of each user experiment, this section attempts to communicate the higher-order descriptions of each Experience. For this conversation, it is helpful to refer back to Milgram's Reality-Virtuality Continuum (RVC), as was referenced in Chapter 2 of this paper. The RVC helps to categorize the 13 Experiences into six general categories that are more easily discussed:

- Spatial AR
- Optical See-through AR (oAR)
- Video See-through AR (vAR)
- Semi-immersive VR
- Immersive VR; low-fidelity (VRL)
- Immersive VR; high-fidelity (VRH)

While definitions for these six categories are well established in the academic literature (R. Azuma et al., 2001; Bimber & Raskar, 2005; Kato & Billinghurst, 1999; Milgram & Kishino, 1994; Milgram et al., 1995), in relation to this effort, the following Figure was created by the research team to better describe the organization of the numbered Experiences within each of the six categories and place them along the RVC:

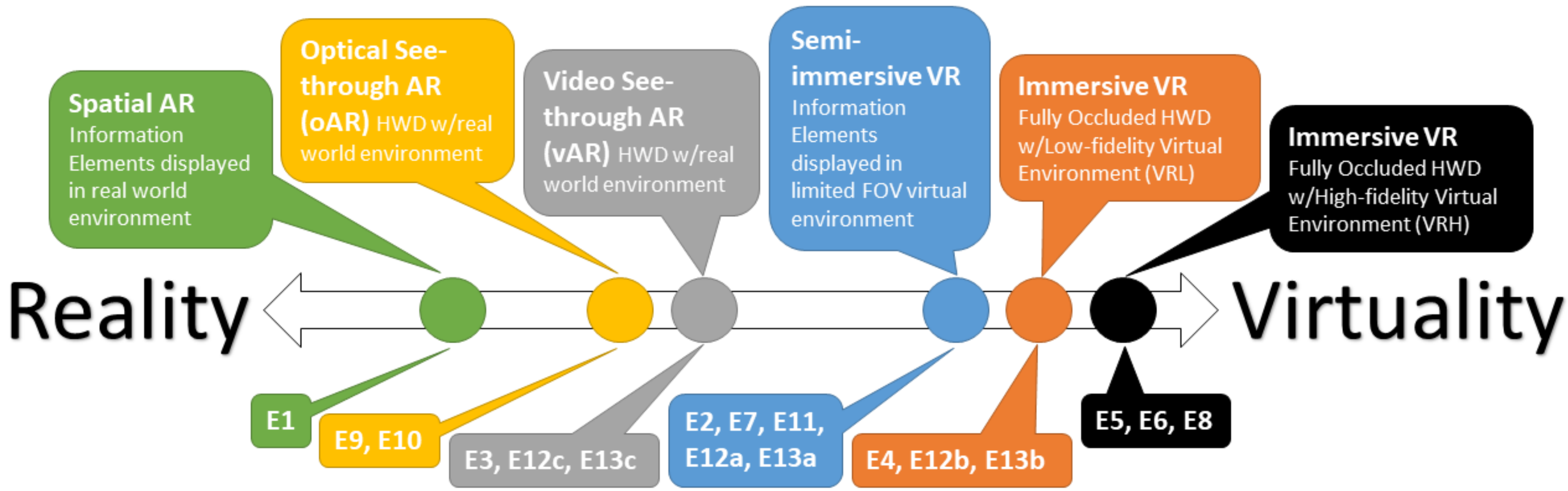


Figure 67: Reality-Virtuality Continuum (RVC)

Based on the previous Figure, the following definitions are therefore provided to describe the difference between all six of the experimentation categories:

- Spatial AR (SAR)
  - Information elements are displayed in the real world envrionment (e.g., a projector places digital infomration elements on real world surfaces)
- Optical See-through AR (oAR)
  - Optical see-through AR HWD with a real world environment viewed by the user (e.g., a traditional AR HWD that projects information elements into the eyes of a user and also allows the user to view the real world)
- Video See-through AR (vAR)
  - Video see-through AR HWD with the real world envrionment viewed by the user on a digital screen (e.g., camera pass-through mode on an Android device where the information elements and real world envrionment are both projected into the eyes of the user)
- Semi-immersive VR (SiVR)
  - Information elements are displayed in a limited FOV virtual environment (e.g., a simulated environment is projected on a single wall while a HWD projects information elements into the eyes of the user)
- Immersive VR; Low-fidelity (VRL)
  - Fully occluded HWD with a low-fidelity virtual environment (i.e., a traditional VR headset with a low-fidelity virtual environment that also displays information elements to the user)
- Immersive VR; High-fidelity (VRH)
  - Fully occluded HWD with a high-fidelity virtual environment (i.e., a traditional VR headset that displays a high-fidelity virtual environment and information elements to the user)

With these definitions explained, it is helpful to reflect back on the resutls of the overall research effort and perform an analysis of the lessons learned from each category of Experience on the

RVC.

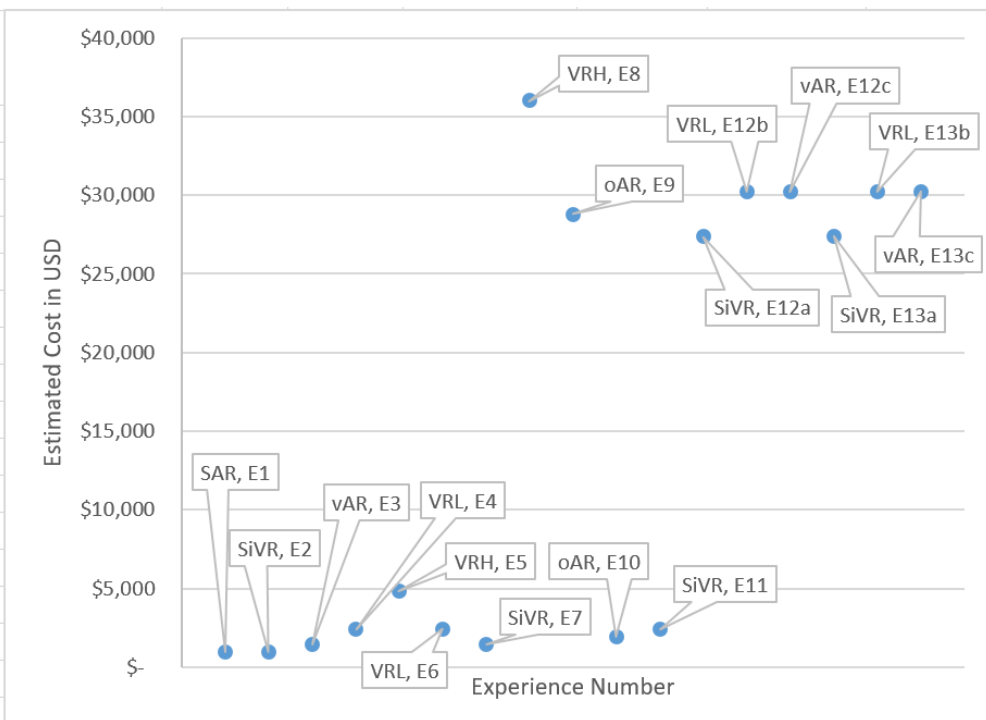
Using this newly described categorical organization along the RVC, with the accompanying color representations, Table 12 can now be represented as follows:

Table 13: Abbreviated Experience Characteristics, sorted Chronologically, with RVC Category

					Interface						
					Interaction				Immersion		
	Name	Description	RVC Category	Costs	User Input	System Output	Software	Functional	Tactical	Spatial	Display Hardware
Experience	E1	Projected elements w/real world	SAR	$ 960					X		Projector
	E2	Projected elements on a projected environment v1	SiVR	$ 960		X		X	X	X	Projector
	E3	Mobile phone w/real world	vAR	$ 1,440		X	X	X	X	X	Mobile Phone
	E4	Mobile phone w/low-fidelity VR world	VRL	$ 2,400	X	X	X	X	X		Mobile Phone
	E5	VR HWD w/high-fidelity VR world	VRH	$ 4,800	X	X	X	X	X	X	VR Headset
	E6	Tethered AR HWD w/low-fidelity VR world	VRH	$ 2,400	X	X	X	X	X		AR Headset
	E7	Tethered AR HWD w/projected environment v1	SiVR	$ 1,440		X		X	X	X	AR Headset
	E8	Tethered AR HWD w/high-fidelity VR world	VRH	$36,000	X	X	X	X		X	AR Headset
	E9	Tethered AR HWD w/real world v1	oAR	$28,800	X	X	X	X		X	AR Headset
	E10	Large FOV AR HWD w/real world	oAR	$ 1,920	X	X	X	X	X	X	AR Headset
	E11	Projected elements on a projected environment v2	SiVR	$ 2,400	X	X		X	X	X	Projector
	E12a	Untethered AR HWD w/projected environment v2	SiVR	$27,360	X	X	X	X	X	X	AR Headset
	E12b	Mobile phone w/low-fidelity VR world v2	VRL	$30,240	X	X	X	X	X	X	Mobile Phone
	E12c	Mobile phone w/real world v2	vAR	$30,240	X	X	X	X	X	X	Mobile Phone
	E13a	Untethered AR HWD w/projected environment v3	SiVR	$27,360	X	X	X	X	X	X	AR Headset
	E13b	Mobile phone w/low-fidelity VR world v3	VRL	$30,240	X	X	X	X	X	X	Mobile Phone
	E13c	Mobile phone w/real world v3	vAR	$30,240	X	X	X	X	X	X	Mobile Phone

6.3: Financial Investment

One of the most unique contributions of this research effort are the provided development costs of each Experience. Though it is challenging to compare work output amongst organizations, due to the complexity of the human beings that make up organizations, quite frankly, the 13 Experiences described in this effort represent a large enough sample of work in sufficient detail, that future researchers should be able to replicate a given Experience within a reasonable margin of financial certainty. More accurately, one would expect that the lessons learned throughout this effort that are being shared in this paper would better prepare future research opportunities so that less errors would be experienced by other research teams, thereby resulting in somewhat reduced Experience costs for the same execution path. When chronologically ordered, the following Figure shows a more visual cost of each Experience during this effort:



Name	RVC Category
E1	SAR
E2	SiVR
E3	vAR
E4	VRL
E5	VRH
E6	VRL
E7	SiVR
E8	VRH
E9	oAR
E10	oAR
E11	SiVR
E12a	SiVR
E12b	VRL
E12c	vAR
E13a	SiVR
E13b	VRL
E13c	vAR

Figure 68: Total Experience Cost, Chronologically Ordered

One major finding is displayed quite poignantly by this Figure; the large financial gulf between the Experiences that cost less than $5,000 and the ones that cost over $27,000. When this effort was being explored, there was a perception within the research team that there was a gradual increase in the actual costs of each Experience as features were explored and experimented with. While the financial expenses of the overall effort were indeed gradually increasing year over year, the post hoc analysis of each individual Experience cost was not as linear. The advanced software programming required to build the most expensive Experiences meant a significant financial investment was required to build those prototypes. Additionally, it is important to note that the more expensive Experiences did not fall into any single RVC category; they spanned all but the SAR option. And as a result, every RVC category had an available option within the cheaper $5K or less investment bracket.

When reflecting on the lack of useful user data results in E8 and E9, it is difficult to justify such costs as simply a lesson learned. Experiences in the lower bracket of less than $5,000 were of sufficient interface and interaction fidelity that users perceived them to be “realistic enough” for the purposes of experimentation. They expressed being able to sufficiently feel as though the Experience were a possible future in which they would perform their work-related duties. This, in turn, meant that gathering useful feedback was possible and that such feedback did, indeed, inform future rapid design and prototype iterations throughout the research effort. The only mechanism that was gained from the higher bracket of over $27,000 was a higher fidelity version of interaction. Instead of using “Wizard of Oz”-type user input and system output triggers that were controlled by the research team itself, the $22,000+ investment in programmer skills created Experiences which were more software-based and could be controlled in their entirety by the first responder. While this higher fidelity of interaction was valuable, it is postulated that such an extreme investment is only justified in a final or near-final prototype version. When other less costly measures are taken (e.g., PD-based storytelling, immersive virtual video game environments) to create an Experience, very little UCD-based feedback is missed.

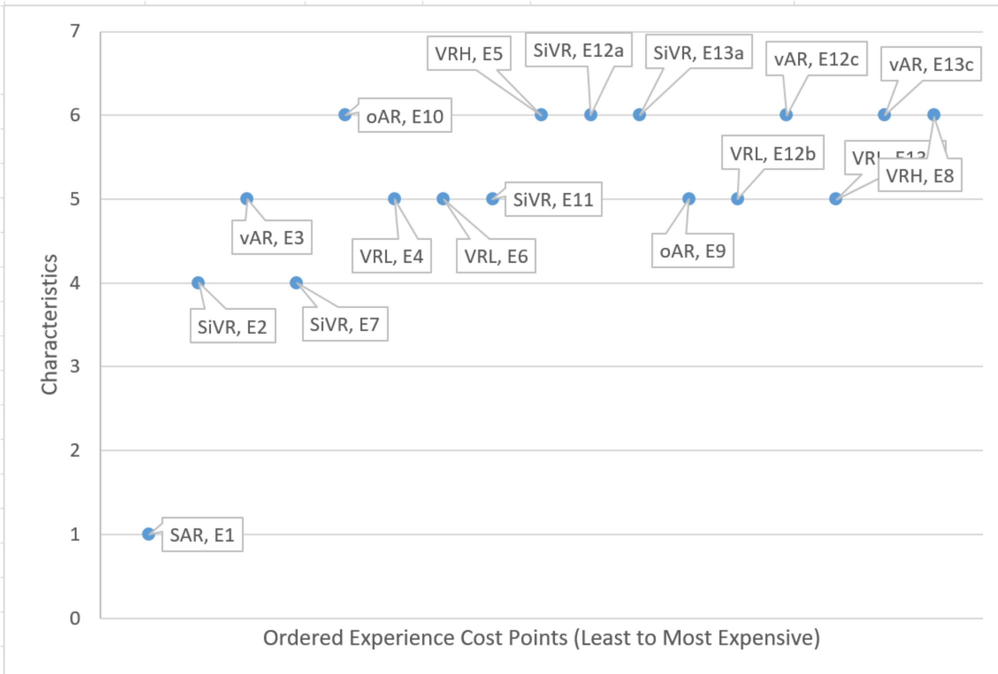
Additionally, in terms of the Experiences of this study, there are myriad Experience options within the lower $5,000 or less investment option. No single Experience was the sole repository of useful UCD-based feedback during the research effort. Depending on the intended results of the

Experience, all of the low-cost options are capable of addressing each of the characteristics of interface, interaction, and immersion when carefully planned and executed. In E10, for example, the details of the environment (see A9 for more detail) in which the first responders were placed were paramount to creating a setting in which users felt the Experience was an accurate depiction of a real futuristic life. As mentioned previously, the first responder participant did not actually control any of the interaction systems, but the Experience included an appropriate level of detail that it was successful in providing the feeling that it was realistic.

Lastly, there do not appear to be any obvious Experience choices in terms of financial investment between $5,000-$27,000, which is a surprising result to the research team. Even during Experiences that were very limited in task scope (e.g., E12 and E13), the investment in financial resources to create the MR HWD software system was over $27,000. The research team expected that there would be a gradual progression in the cost of developing more intricate and capable Experiences as the effort proceeded, but this was not the case for this research. It was also expected that certain characteristics of each Experience would not be possible to communicate at lower costs (e.g., tactical immersion, software interaction). While it is true that entry-level fidelities of all characteristics are possible at lower-cost Experiences, a fully interactive system is likely to continue to require significant financial investment in MR systems.

6.4: Characteristic Investment

While the previous tables and chart have described the financial investments required to develop each Experience of the research effort, in chronological order, it is suggested that the Experiences also be analyzed through a post hoc lens. By determining the tangible inputs and outputs of UCD-based 3D MR HWD rapid design and prototype iterations, the research team can determine where the intersection of financial investment and other metrics might provide insight to future experiments. One possible nexus of data is when each Experience is analyzed through the total number of characteristics that are employed within it. One could easily justify that an Experience with more characteristics is more valuable when UCD-based feedback is sought out. Next, the Table of Experience Characteristics is organized by price, from lowest to highest. A new column is also added that shows how many total characteristics are included in each Experience. When the total number of Experience characteristics is organized by price, the following Figure is made:



Name	RVC Category
E1	SAR
E2	SiVR
E3	vAR
E4	VRL
E5	VRH
E6	VRL
E7	SiVR
E8	VRH
E9	oAR
E10	oAR
E11	SiVR
E12a	SiVR
E12b	VRL
E12c	vAR
E13a	SiVR
E13b	VRL
E13c	vAR

Figure 69: Number of Experience Characteristics, Organized by Total Price

With a range of zero to six possibilities, this Figure shows that while E1 only included a tactical characteristic of the immersion category, the remaining Experiences include four to six characteristics each. In terms of the characteristics of immersion, user input becomes an established standard by E8, system output by E2, software by E12. In terms of the interface category, the interaction characteristics are established by E12 and the functional characteristic by E3. Tactical immersion is established by E10, but spatial immersion continues to fluctuate until E13c. All characteristics are included in E5, E10, E12a, E12c E13a, and E13c. This comparison of price to characteristics seems to indicate that the Experiences to the right of E5 are too expensive when compared to their relative UCD-based feedback because they do not include an increased number of Experience characteristics. In fact, E8 and E9 include only five characteristics and entirely failed to execute on tactical immersion because the Experiences were unreliable and did not respond rapidly or reliably.

The less investment-heavy Experiences are a result of being able to employ lower-fidelity interface, interaction, and immersion characteristics in the lower-investment bracket of less than $5,000. While increased levels of fidelity in these characteristics were possible at the higher investment levels, they were not always guaranteed. For example, E10 was developed to showcase a large FOV AR HWD. While it *simulated* the interface, interaction, and immersion categories of characteristics well, it was a tethered Experience that did not allow the user any customization control with their HUD. Users consistently provided feedback that a more realistic MR HWD system for future first responders would require an untethered experience and first-hand interaction control (at an increased level of fidelity) in order to provide increased levels of immersion overall. E12 and E13 were the only Experiences to successfully gather feedback on how users could customize and control the information displayed on their futuristic MR HUD.

It is valid to question the usefulness of E5, E8, and E9 due to the large investment in development required during each prototype build and the very low level of UCD-based feedback gathered during experimentation. E5 was only briefly useful to the research effort because a large section of first responder participants suffered from virtual simulation sickness. This reduced the amount of time that participants were able to use E5 overall, which also reduced the amount of UCD-based feedback that could be gathered. By decreasing the pool of participants, it was harder for the research team to find consensus amongst the remaining first responders. Additionally, this meant that in order to gather the necessary feedback to move forward in the effort, the research team was required to employ a longer-form semi-structured interview with those participants that were unaffected by the virtual simulation. And because a participant's ability to remain unaffected by the simulation was a completely random event, no preparation could be made before an interview in the case that the simulation sickness had to be addressed. For Experiences after E5, a completely virtualized environment could never again be the only instantiation for an experiment condition (e.g., E12 and E13 had VR versions for participants that desired to try that simulation, but also AR versions for all participants to test).

6.5: Recommended Path

The next Table shows the final data sorting arrangement for this chapter. A slight variation of the previous Table, this one sorts the entire Experience list by the number of total characteristics demonstrated in each experiment, then sorts the Experiences within that number of characteristics by ascending price. In other words, this Table demonstrates the cheapest and most expensive Experience options within a specific number of total characteristics:

Table 14: Abbreviated Experience Characteristics, Sorted by Total Characteristics, then Price

					Interface				Immersion			
					Interaction							
	Name	Description	RVC Category	Costs	User Input	System Output	Software	Functional	Tactical	Spatial	Display Hardware	Total Characteristics
Experience	E1	Projected elements w/real world	SAR	$ 960					X		Projector	1
	E2	Projected elements on a projected environment v1	SiVR	$ 960		X		X	X	X	Projector	4
	E7	Tethered AR HWD w/projected environment v1	SiVR	$ 1,440		X		X	X	X	AR Headset	4
	E3	Mobile phone w/real world	vAR	$ 1,440		X	X	X	X	X	Mobile Phone	5
	E4	Mobile phone w/low-fidelity VR world	VRL	$ 2,400	X	X	X	X	X		Mobile Phone	5
	E6	Tethered AR HWD w/low-fidelity VR world	VRL	$ 2,400	X	X	X	X	X		AR Headset	5
	E11	Projected elements on a projected environment v2	SiVR	$ 2,400	X	X		X	X	X	Projector	5
	E9	Tethered AR HWD w/real world v1	oAR	$28,800	X	X	X	X		X	AR Headset	5
	E12b	Mobile phone w/low-fidelity VR world v2	VRL	$30,240	X	X	X	X	X		Mobile Phone	5
	E13b	Mobile phone w/low-fidelity VR world v3	VRL	$30,240	X	X	X	X	X		Mobile Phone	5
	E10	Large FOV AR HWD w/real world	oAR	$ 1,920	X	X	X	X	X	X	AR Headset	6
	E5	VR HWD w/high-fidelity VR world	VRH	$ 4,800	X	X	X	X	X	X	VR Headset	6
	E12a	Untethered AR HWD w/projected environment v2	SiVR	$27,360	X	X	X	X	X	X	AR Headset	6
	E13a	Untethered AR HWD w/projected environment v3	SiVR	$27,360	X	X	X	X	X	X	AR Headset	6
	E12c	Mobile phone w/real world v2	vAR	$30,240	X	X	X	X	X	X	Mobile Phone	6
	E13c	Mobile phone w/real world v3	vAR	$30,240	X	X	X	X	X	X	Mobile Phone	6
	E8	Tethered AR HWD w/high-fidelity VR world	VRH	$36,000	X	X	X	X	X	X	AR Headset	6

Additional highlighting was added to this Table in order to emphasize the two cost segments of the Experiences: less than $5,000 and more than $27,000. This ordering should provide two potential development paths for future research that logically progress according to cost and the number of characteristics of the Experience. The following Figure shows the development ordering for Experiences that require less than $5,000 in investment each:

Name	RVC Category
E1	SAR
E2	SiVR
E3	vAR
E4	VRL
E5	VRH
E6	VRL
E7	SiVR
E8	VRH
E9	oAR
E10	oAR
E11	SiVR
E12a	SiVR
E12b	VRL
E12c	vAR
E13a	SiVR
E13b	VRL
E13c	vAR

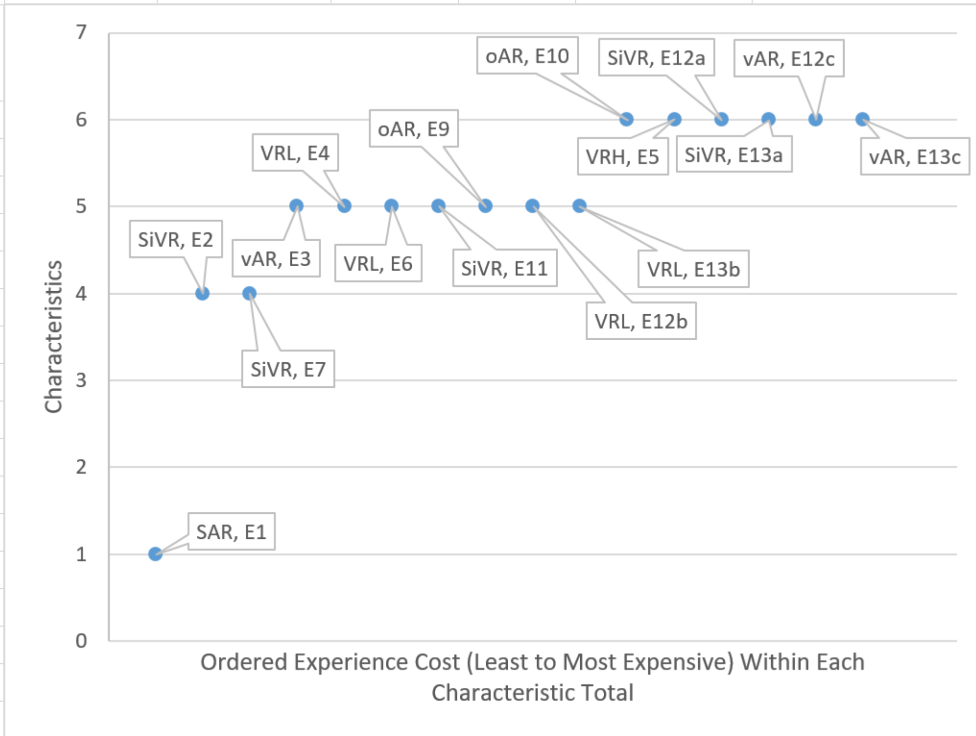
Figure 70: Recommended Development Path, Prioritized by Total Price Within Each Total Category, <$5,000 Investment Each

Moving from left to right, this Figure indicates the following categorical segmentation within the given constraints:

- One Characteristic:
  - E1
- Four Characteristics:
  - E2, E7
- Five Characteristics:

    - E3, E4, E6, E11
- Six Characteristics:
    - E10, E5

By referring to this Figure and the previous Table, future researchers can choose which characteristics they want to focus on in their own efforts and confidently build their own rapid prototype version of the described Experience in order to gather useful UCD-based feedback. With the gift of hindsight, the research team suggests employing E1, E2, E7, E11, then E10 as the recommended path of development for less than $5,000 each. Of course this path does not include any additional hardware investment that would be required to create each Experience (e.g., a mobile phone is a hardware device that most developers own themselves, while a large FOV AR HWD is something very few researchers currently posess), but each of these experiments employed a sufficient number of characteristics that could be demonstrated at a low financial investment cost in order to gather useful user-centered feedback on MR HWD systems. Additionally, it is important to note that there are different options that can be deployed across the RVC categories. Although SAR was only present during E1, the other RVC categories span the number of characteristics. For instance, the VRH category option was deployed with five and six characteristic versions. If the research budget is less than $40,000, then the remainder of Experiences from this effort can also each be considered:



Name	RVC Category
E1	SAR
E2	SiVR
E3	vAR
E4	VRL
E5	VRH
E6	VRL
E7	SiVR
E8	VRH
E9	oAR
E10	oAR
E11	SiVR
E12a	SiVR
E12b	VRL
E12c	vAR
E13a	SiVR
E13b	VRL
E13c	vAR

Figure 71: Recommended Development Path, Prioritized by Total Price Within Each Total Category, <$40,000 Investment Each

For the interest of the reader, the estimated total cost to develop all of the Experiences in this research effort is approximately $260,000. Moving from left to right again, the critical path of development where investment cost is the least constrained variable, is enumerated next. Differences from the previous list are in **bold**:

- One Characteristic:
  - E1
- Four Characteristics:
  - E2, E7
- Five Characteristics:
  - E3, E4, E6, E11, **E9, E8**
- Six Characteristics:
  - E10, E5, **E12, E13**

The recommended development path, given the benefit of hindsight to the research team at this higher financial investment level, remains largely the same. It is suggested that mobile-phone-based MR HWDs and VR-based Experiences be avoided, but the following options be explored in this order: E1, E2, E7, E11, E10, E12a, E13a. Only E12a and E13a involve a high level of financial investment in order to create them.

While E10 and E5 appear to be the best value in terms of the most characteristics for the least financial investment cost, when analyzed through the lens of gathering useful UCD-based feedback that was focused on the first responder assault role in the common response scenario, E10 was substantially more effective than E5. As previously expressed, significant physiological difficulties were experienced with completely virtualized testing environments, therefore it was much more difficult to successfully elicit appropriate user-centered feedback during VR-based Experiences. While both of these Experiences were fully simulated in terms of system output and involved little customization of the futuristic HUD, they were very effective in placing the participant in the right frame of mind to gather useful feedback. And although both of these Experiences were also tethered in place, there was enough tactical and spatial immersion that participants often overlooked the stationary nature of E10 and E5, therefore in cases where the environmental immersion of an experience can be enhanced at a low cost, especially as related to spatial and tactical immersion, it is essential that appropriate PD-based storytelling methods be coupled with the immersive situation in order to place the participant in the most tangibly realistic scenario possible. In cases where little or no immersion can be provided, the research team found that the only feedback that is gathered from participants is related to the lack of experiential immersion, which proves less-than-effective in making useful progress in the overall research effort.

Perhaps more interesting to academic researchers should be the placement of E11 in the recommended path. While also stationary in terms of physical location, the tactical and spatial immersion characteristics of the projector display hardware, utilizing a video game environment for contextual information, proved to be highly engaging to the user, which made the data gathering process highly effective to the research team. Hundreds of data points were logged in an otherwise simplistic research testbed that generally eliminated the technical shortcomings of previous Experiences.

When this same video game-type immersive environment was extended to E12 and E13, combined with a user-customizable futuristic HUD, each participant was even more engaged in the overall Experience, providing a perception that such a technology would actually be feasible one day. Throughout all of the Experiences, a few key findings were found to apply across the research effort:

- User's expect even low-fidelity prototypes to have advanced capabilities in terms of interface, interaction, and immersion
- User's often require first-hand experience of a prototype in order to provide informed feedback; they want to try it out themselves
- Storyboarding is critical to visualizing and discussing the details of a futuristic research effort
- PD-based storytelling is critical to providing the appropriate context and the mental models of a futuristic research effort
- Semi-structured interviews were highly effective in gathering UCD-based feedback
- Gesture elicitation was highly effective in gathering feedback on futuristic actions to perform in a 3D MR HWD system.

6.6: RVC Impact

As has been described previously, an analysis of the impact of the six RVC categories on this research effort is in order. The definitions of each category have already been explained, but the follow subsections provide a transcendent level of discussion regarding each RVC option.

6.6.1: SAR

Although a SAR experiment was always the intended target of the entire research effort, only one Experience fell into this category along the VR scale. This was largely due to technical limitations of the hardware devices used in early experiments, but could now be tested in future Experiences with currently available consumer devices. In fact, had more resources been available to the research team, this would have been the next logical step in the experimentation progress of the common response scenario with first responders. As was discussed during O2, E1 was the first experiment with first responders and although brief in its duration and extremely basic in its execution, it was highly effective at gathering useful UCD feedback from first responders at a very low cost. In fact, the visual graphics created during the SAR experiment could largely be considered the same information elements that were only subtly refined during every subsequent experiment and were very close to matching that basic training HUD that was developed much later on in the research effort. The results of the SAR experiment were almost prophetic to the overall research effort.

6.6.2: oAR

Two Experiences fall into the oAR category: one in each of the two price brackets, but the cheaper option was much more effective in gathering useful UCD feedback because it included tactical immersion as its sixth characteristic. Without positive tactical immersion, no category along the RVC proved worth the investment to create that Experience. The addition of a highly complex hardware and software system ultimately made the more expensive oAR Experience fail in every experiment with every participant. Going forward, it is recommended that even with professional software developers, that only small and incremental steps be taken with each iteration of each Experience. The research team simply attempted to create a feature set that was comprehensive, but quite unstable.

6.6.3: vAR

Three Experiences fall into the vAR category: one in the cheaper price bracket and the remaining two options in the higher price bracket. Each of these experiments was performed with an Android-based mobile phone, which provided some benefits (e.g., the ability to quickly create a simple application for the cheaper vAR Experience, quick deployment of the software to myriad hardware devices), but resulted in some drawbacks that were not to be ignored. With all vAR options, the FOV of the Experience was completely dependent on the hardware device and was almost always smaller in size than any dedicated consumer AR HWD. This provided negative feedback from most first responders because they sought ever more realistic experimental testbeds as the research effort progressed. Additionally, the vAR Experience did not include a stereoscopic view of the world, meaning most participants could not easily adapt to a simple video camera style feed of the real world displayed to them. It was also difficult for first responders to use any interaction and customization devices in any experiment because of the unnatural display perspective of mobile phone devices. In all cases, the research team suggests that, for this domain of application, the vAR category of Experience be avoided.

6.6.4: SiVR

Five Experiences fall into the SiVR category: three in the cheaper price bracket and two in the upper price bracket. For the purposes of this study, SiVR was by far the most effective set of experiments in gathering useful UCD feedback from first responder users in the common response scenario. It was quickly discovered by the research team that the mental and physical immersion of the first responder into the common response scenario was the single most important category of measure. Tactical and spatial immersion could not be ignored. When these characteristics were ignored in other Experiences, the participant was often distracted from the intended purpose of the experiment and did not provide helpful feedback to progress the effort to the next step. With a choice of several financial commitment options as well, future research can employ Wizard of Oz-type SiVR scenarios in a cost effective manner before moving on to more high-fidelity and complex software system rapid prototypes. The effectiveness of SiVR is why all five of the Experiences within this RVC category are included in the recommended path; no other RVC category holds this distinction.

6.6.5: VRL

Four Experiences fall into the VRL category: half of them in the less expensive investment bracket and the other half in the more expensive bracket. While each of these prototype systems had great promise, the prevalence of virtual reality simulation sickness could not be ignored. Each of these experiments included functional user input that was reliable and provided the participant a level of customization and interaction that was not present in most Experiences. Unfortunately, all VRL experiments lacked any spatial immersion, which led to a lack of overall immersion that was distracting to participants and reduced the level of useful UCD feedback that could be gathered by the research team. And although a higher fidelity of user customization was available in each VRL Experience, because the participant was not able to physically see and reference their hands in relation to the interaction devices that were available to them, most were unable to use those devices without a level of training that was inconvenient and less-than-helpful to a rapid iteration prototype environment. It is suggested that future researchers do not explore VRL experimental testbeds in this domain.

6.6.6: VRH

Only two Experiences fall into the VRH category; one in each price bracket. Similarly to VRL, they included user customization features, but because users were unable to see the interaction devices, they were unable to use them. Simulation sickness also plagued this experimental category. Spatial immersion was high in these Experiences, but with fewer first responder participants to interview and both VRH experiments being the highest options in their categories, the research team suggests that VRH prototypes be avoided in this domain and SiVR, oAR, or SAR rapid prototype options be explored in their place.

6.7: Research Answers

- *RQ1: What information do first responders expect to be available to them in a futuristic mobile HWD MR interface?*

A1 and A2 utilized traditional methods from user-centered design (UCD) practices, including semi-structured interviews, PD-based storytelling, brainstorming, card sorting, affinity diagrams, and storyboarding to examine RQ1. The research team created an exhaustive list of information elements that first responders expected to be available to them in a futuristic mobile head-word MR interface (i.e., the list of 81 information elements), as was shown in Table 4: Information Element List from A2.

- *RQ2: What information is most critical to the first responder to allow them to perform their job safely?*

A3 and A4 developed a more targeted version of the information elements from A2. Using a card sorting method and storyboard format, which incorporated five specific time points, the research team successfully developed a standard training HUD that would communicate the most critical information quickly to the user. This training HUD was initially described in Table 10: Common Critical Elements, Arranged by Time Point, AKA, the “Training HUD”. Although iteratively refined during each subsequent Activity, the final visual representation of the training HUD was presented in Figure 63: E13a Training HUD.

- *RQ3: How can critical interface information be communicated to first responders utilizing multiple modalities (e.g., visual, aural, haptic) of notification?*

A5 provided the initial conceptual designs to answer RQ3. These concepts were experimented with during subsequent Activities throughout this research effort, but were largely addressed in A6 and A7.

- *RQ4: How does the context in which a user experiences a prototype effect the interface feedback they provide?*

E1-E7 provided a rich breadth of prototype experiences, which incorporated different combinations of interface, interaction, immersion, and display hardware. This RQ was continually explored throughout the rest of the research effort and direct analyses amongst the Experiences will be described in more detail later in this chapter.

- *RQ5: How do first responders desire to interact with critical information?*

Although initial feedback on the topic of user preference in interaction had been gathered informally during O1 and O2, A8 marked the introduction of fully developed software-based interaction prototype Experiences. E8 and E9 implemented user-based interaction motifs formally into the experimentation testbed while A9 and A10 refined the interaction concepts into their final iterations.

- *RQ6: How does the context in which a user experiences a prototype effect the interaction* *feedback they provide?*

Throughout the entire research effort, differing fidelities of interface, interaction, and immersion were experimented with in order to gather useful UCD-based feedback on various aspects of 3D MR HWD development. RQ6 addresses these changes as they relate to interaction specifically and were addressed in each Activity, with E13 representing the final iteration of the research efffort's focus on user interaction.

6.8: Limitations of the Work

The most common potential research shortcoming that arose during this effort was an unbalanced weight of external stakeholders (e.g., no formal representation of top-level command-based leadership at a county/state/federal rank) to internal first responder team-based SMEs (e.g., those who respond to emergency incidents and their direct supervisors). Those included in the ideal first responder user profile were generally at the first level of management (e.g., Lieutenant, Seargent) depending on local rank hierarchies, so they were intimately familiar with what each team member was doing under their leadership. And while this still includes the approximately one standard deviation representation of the number of persons involved within the first responder community, the importance of higher-level management and command structures is important to other stakeholders and cannot be ignored when using a larger perspective of more large-scale first responder scenarios (e.g., devastation from natural disasters, states of emergency). While an attempt at a more holistic representation was more accurately included in the later experiments of this research effort, they should be considered earlier in future exploratory work in this domain so that researchers can gather first-hand experiential data those participants that specialize in that level of detail instead of second and third-hand recitations of perceived user-centered informational requirements from lower ranking internal team-based SMEs. While there was a need to focus on a single ideal user profile initially in order to clearly focus the effort, including in-depth perspectives for additional levels of understanding and more macro-level operations could potentially change the findings of this research effort.

The first iteration of the card sorting experiment of A3 was significantly less-effective than the research team had hoped it to be. Without the time point context explored in A4, A3 resulted in inconsistent results toward the progress of O1. Future research should ensure that more diligent attempts are made to fully understand the context of the experiential testbed by asking more questions of the end-user participants. This increased level of detail is what was required to make the second iteration of the card sorting method in A4 a success.

Until the end of A7, the focus of the research effort was solely on a single user in a single role with a singular job to accomplish in a singular context. While the context of the common response scenario was eventually deemed to be much more generalizable than initially described, the groups of users interviewed during A1-A8 could have unfairly biased the results of the work provided to the more generalized population of first responders in A9 and A10. The more general population was asked for feedback on the established research artifacts (e.g., team member roles, ideal user profile, critical information elements, common response scenario, representations of individual information elements), but very little critique was given to these items.

Significant time and resources were wasted on E8 and E9. It is recommended to future researchers that such a large-scale Experience be avoided and more specifically targeted Experiences be developed instead, like the examples given in E12 and E13.

Strategic and narrative characteristics did not vary in any Experience of this effort. Future research should explore these two important aspects of immersion and their effect on gathering useful UCD-based feedback. Certainly E12 or E13 could now be easily modified to address strategic and narrative changes and to observe the changes in user behavior.

Most of the data gathered and analyzed in this effort incorporated qualitative feedback. While this is a comfortable place of operation within the UCD community, as has been discussed in Chapter 2, related scientific fields might dismiss the contributions of this effort due to the abundance of qualitative information. Future research should not reduce the amount of qualitative data gathered, but should include more quantitative numbers as well.

Because of the nature of this rapid design and prototype effort, many of the constraints of this research are direct results of the hardware and software systems that were available to and utilized by the research team. As time goes on, it is expected that actual costs for developing a specific Experience will continue to decrease as the tools available become more powerful to software developers and more accessible to the general population.

6.9: Future Research

As has been suggested, the research team expects that the detail of lessons learned that have been described throughout this paper will help future experiments to be more effective in their execution and less costly in their required financial investment. The methods-based approach to each Experience should provide the reader with the appropriate guidance to implement whichever experiment characteristics they desire to test from the interface, interaction, and immersion categories.

A series of experiments that utilize the recommended path should be explored to determine whether the results of this effort can be replicated elsewhere in either the first responder domain or related domains. The results of those experiments should be analyzed and added to the experience characteristic table shared herein.

The more effective methods used in this effort should be applied to other domains of interest in MR. Action elicitation experiments should be performed to determine whether the results from A9 are applicable to the general population. These actions would serve to help establish standards for interacting with technology in a 3D MR environment.

The information element interface representations tested throughout this effort should also be explored in additional domains. This would help establish acceptable standards for what specific pieces of information should look like in a 3D MR environment.

It is also suggested that future research opportunities actively seek middle-of-the-range investment opportunities for specific Expereinces. At noted previously, no Experience cost between $5,000-$27,000. It would be interesting to develop an Experience that cost $10,000, $15,000, and $20,000 in order to perform an analysis of the characteristics included in that range as compared to the results described during this research effort.

Team-based Experiences were never addressed during this effort. While they were conceptually outlined in storyboard format, future design and rapid prototype efforts could include multi-perspective and multiple team roles in the common response scenario. New information elements would need to be developed and it is expected that changes to traditional information gathering UCD-based methods might be necessary to make these higher-level experiments successful.

Finally, it is proposed that after following the recommended path proposed in this research, that the UX testbed be taken out into more real world environments and experiments be run in a non-traditional laboratory setting. Technical limitations of HWD systems will need to be considered (such as the inability to see AR objects in a brightly lit room or an outside environment) in these future designs, but these more realistic types of scenarios should be explored in future work.

7: References

A Survey of Augmented Reality Technologies, Applications and Limitations. (2010). *International Journal of Virtual Reality*, *9*(2), 1–19.

Achen, C. H., & Snidal, D. (1989). Rational Deterrence Theory and Comparative Case Studies. *World Politics*, *41*(2), 143–169. https://doi.org/10.2307/2010405

Adams, A., Soumerai, S. B., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *International Journal for Quality in Health Care*, *11*(3), 187–192.

Adams, E. (2004, July 9). The Designer's Notebook: Postmodernism and the 3 Types of Immersion. Retrieved December 1, 2017, from https://www.gamasutra.com/view/feature/130531/the_designers_notebook_.php

Alghamdi, A. (2010). Remote Access to Employee Data in the Saudi Royal Navy. Retrieved from http://www.macs.hw.ac.uk/~pjbk/projectdata/archive/2011/pg/aaa148_full_text.pdf

Altman, D. G., & Bland, J. M. (2011). Brackets (parentheses) in formulas. *BMJ*, *343*, d570. https://doi.org/10.1136/bmj.d570

Altman, D. G., & Bland, J. M. (2014a). Uncertainty and sampling error. *BMJ*, *349*, g7064. https://doi.org/10.1136/bmj.g7064

Altman, D. G., & Bland, J. M. (2014b). Uncertainty beyond sampling error. *BMJ*, *349*, g7065. https://doi.org/10.1136/bmj.g7065

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, *21*(6), 34–47. https://doi.org/10.1109/38.963459

Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Baber, C., Sharples, M., Boardman, M., Price, A., & Haniff, D. J. (2001). Requirements Engineering of Personal Technology for Police Officers. In *Interact* (pp. 512–520). Retrieved from https://pdfs.semanticscholar.org/73de/f7d215d2145fd22f5380822ce3d10283c605.pdf

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2000). Measuring the effects of interruptions on task performance in the user interface. In *2000 IEEE International Conference on Systems, Man, and Cybernetics* (Vol. 2, pp. 757–762 vol.2). https://doi.org/10.1109/ICSMC.2000.885940

Baxter, K., Courage, C., & Caine, K. (2015). *Understanding Your Users: A Practical Guide to User Research Methods*. Morgan Kaufmann. Retrieved from https://books.google.com/books?hl=en&lr=&id=l9-cBAAAQBAJ&oi=fnd&pg=PP1&dq=user+experience+research+methods&ots=dBLIRzUpTZ&sig=JoRm5KAkUMmcX4yQzYcNnesLxZw

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, *11*(3), 369–386. https://doi.org/10.2307/248684

Berg, B. L. (1998). *Qualitative research methods for the social sciences* (3rd ed). Boston: Allyn and Bacon.

Billinghurst, M., Clark, A., & Lee, G. (2015). A Survey of Augmented Reality. *Foundations and Trends® in Human–Computer Interaction*, *8*(2–3), 73–272. https://doi.org/10.1561/1100000049

Bimber, O., & Raskar, R. (2005). *Spatial Augmented Reality: Merging Real and Virtual Worlds*. CRC Press.

Bjork, S., & Holopainen, J. (2004). *Patterns in Game Design*. Hingham, UNITED STATES: Charles River Media. Retrieved from http://ebookcentral.proquest.com/lib/vt/detail.action?docID=3135809

Bland, J. M., & Altman, D. G. (1996). Statistics Notes: Measurement error. *BMJ*, *313*(7059), 744. https://doi.org/10.1136/bmj.313.7059.744

Bland, J. M., & Altman, D. G. (2015). Statistics Notes: Bootstrap resampling methods. *BMJ*, *350*, h2622. https://doi.org/10.1136/bmj.h2622

Blomberg, J., & Burrell, M. (2009). An Ethnographic Approach to Design. In *Human-Computer Interaction. Development Process* (Second, pp. 71–94). CRC Press.

Bødker, K., Kensing, F., & Simonsen, J. (2010). Participatory Design in Information Systems Development. In *Reframing Humans in Information Systems Development* (pp. 115–134). Springer, London. https://doi.org/10.1007/978-1-84996-347-3_7

Bolt, R. A. (1980). “Put-that-there”: Voice and Gesture at the Graphics Interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 262–270). New York, NY, USA: ACM. https://doi.org/10.1145/800250.807503

Bowman, D. A., & Hodges, L. F. (1999). Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, *10*(1), 37–53.

Bowman, D. A., Johnson, D. B., & Hodges, L. F. (2001). Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoperators and Virtual Environments*, *10*(1), 75–95.

Bowman, D. A., Kruijff, E., LaViola Jr, J. J., & Poupyrev, I. (2001). An introduction to 3-D user interface design. *Presence: Teleoperators and Virtual Environments*, *10*(1), 96–108.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer*, *40*(7).

Box, G. E. P., Hunter, J. S., & Hunter, W. G. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery, 2nd Edition* (2nd edition). Hoboken, N.J: Wiley-Interscience.

Bragdon, A., DeLine, R., Hinckley, K., & Morris, M. R. (2011). Code space: touch+ air gesture hybrid interactions for supporting developer meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (pp. 212–221). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2076393

Brener, N. D., Billy, J. O. G., & Grady, W. R. (2003). Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *33*(6), 436–457.

Budwig, M., Jeong, S., & Kelkar, K. (2009). When User Experience Met Agile: A Case Study. In *CHI ’09 Extended Abstracts on Human Factors in Computing Systems* (pp. 3075–3084). New York, NY, USA: ACM. https://doi.org/10.1145/1520340.1520434

Carroll, J. B. 1916-2003 (John B. (1993). *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Catani, M. B., & Biers, D. W. (1998). Usability Evaluation and Prototype Fidelity: Users and

Usability Professionals. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*(19), 1331–1335. https://doi.org/10.1177/154193129804201901

Chi, H.-L., Kang, S.-C., & Wang, X. (2013). Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in Construction*, *33*, 116–122. https://doi.org/10.1016/j.autcon.2012.12.017

Clarivate Analytics. (2017, July 13). Web of Science Citing Articles. Retrieved July 13, 2017, from http://apps.webofknowledge.com.ezproxy.lib.vt.edu/InboundService.do?action=search&exclude=exclude

Cook, R., Science, U. of S. B. D. of C., & Science, U. of W. A. D. of C. (1994). *A toolkit for appraising the long term usability of a text editor*. Citeseer.

Crabtree, A. (1998). Ethnography in participatory design. In *Proceedings of the 1998 Participatory design Conference* (pp. 93–105). Retrieved from http://www.cs.nott.ac.uk/~pszaxc/work/PDC98.pdf

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005* (Vol. 1, pp. 886–893 vol. 1). https://doi.org/10.1109/CVPR.2005.177

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, *17*(2), 245–260.

Doug’s 1968 Demo - Doug Engelbart Institute. (n.d.). Retrieved December 1, 2017, from http://www.dougengelbart.org/firsts/dougs-1968-demo.html

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, *14*(4), 532–550. https://doi.org/10.5465/AMR.1989.4308385

Feiten, B., Wolf, I., Oh, E., Seo, J., & Kim, H.-K. (2005). Audio adaptation according to usage environment and perceptual quality metrics. *IEEE Transactions on Multimedia*, *7*(3), 446–453. https://doi.org/10.1109/TMM.2005.846793

Ferscha, A. (2000). Workspace awareness in mobile virtual teams. In *Proceedings IEEE 9th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2000)* (pp. 272–277). https://doi.org/10.1109/ENABL.2000.883742

Findlater, L., Lee, B., & Wobbrock, J. (2012). Beyond QWERTY: augmenting touch screen keyboards with multi-touch gestures for non-alphanumeric input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2679–2682). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2208660

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, *12*(2), 219–245. https://doi.org/10.1177/1077800405284363

Frédéric Le Play. (1878). *Les ouvriers européens*. A. Mame et fils. Retrieved from http://archive.org/details/lesouvrierseuro02playgoog

Fuchs, H., Livingston, M. A., Raskar, R., Colucci, D., Keller, K., State, A., … Meyer, A. A. (1998). Augmented reality visualization for laparoscopic surgery. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98* (pp. 934–943). Springer, Berlin, Heidelberg. https://doi.org/10.1007/BFb0056282

Gabbard, J. L., Swan, J. E., Zedlitz, J., & Winchester, W. W. (2010). More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces. In *2010 IEEE Virtual Reality Conference (VR)* (pp. 79–86). https://doi.org/10.1109/VR.2010.5444808

Gabbard, Joseph L. (1997, December 3). *A Taxonomy of Usability Characteristics in Virtual Environments* (thesis). Virginia Tech. Retrieved from https://vtechworks.lib.vt.edu/handle/10919/35731

Gerring, J. (2004). What Is a Case Study and What Is It Good for? *American Political Science Review*, *98*(2), 341–354. https://doi.org/10.1017/S0003055404001182

Gesture: Definition of Gesture in Oxford Dictionary (American English) (US). (2016, March 11). Retrieved March 11, 2016, from https://www.oxforddictionaries.com/us/definition/american_english/gesture

Gibet, S., Courty, N., & Kamp, J.-F. (2006). *Gesture in human-computer interaction and simulation: 6th International Gesture Workshop, GW 2005, Berder Island, France, May 18-20, 2005 : revised selected papers* (Vol. 3881.). New York: Springer. Retrieved from http://su8bj7jh4j.search.serialssolutions.com.ezproxy.lib.vt.edu/?sid=sersol&SS_jc=TC0000317953&title=Gesture%20in%20human-

computer%20interaction%20and%20simulation%20%3A%206th%20International%20Gesture%20Workshop%2C%20GW%202005%2C%20Berder%20Island%2C%20France%2C%20May%2018-20%2C%202005%20%3A%20revised%20selected%20papers

Good, P. I., & Hardin, J. W. (2012). *Common Errors in Statistics (and How to Avoid Them)*. John Wiley & Sons.

Hartson, H. R., & Pyla, P. S. (2012). *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Elsevier.

Healy, M. E. (1947). Le Play's Contribution to Sociology: His Method. *American Catholic Sociological Review*, *8*(2), 97–110.

Henderson, S., & Feiner, S. (2011). Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics*, *17*(10), 1355–1368. https://doi.org/10.1109/TVCG.2010.245

Henderson, S. J., & Feiner, S. (2009). Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality* (pp. 135–144). https://doi.org/10.1109/ISMAR.2009.5336486

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, *64*, 79–102. https://doi.org/10.1016/j.ijhcs.2005.06.002

Hueni, A., Nieke, J., Schopfer, J., Kneubuhler, M., & Itten, K. I. (2009). The spectral database SPECCHIO for improved long-term usability and data sharing. *Computers & Geosciences*, *35*, 557–565. https://doi.org/10.1016/j.cageo.2008.03.015

Huizinga, D., & Elliott, D. S. (1986). Reassessing the Reliability and Validity of Self-Report Delinquency Measures. *Journal of Quantitative Criminology*, *2*(4), 293–327.

Imai, T., Takeo, H., Yoshimura, M., Sakata, A., Sakakibarai, N., & Sekine, C. (2010). Improving the usability and learnability of a home electric appliance with a long-term usability study. *Journal of Engineering Design*, *21*, 173–187. https://doi.org/10.1080/09544820903316995

Iqbal, S. T., & Horvitz, E. (2007). Disruption and Recovery of Computing Tasks: Field Study, Analysis, and Directions. In *Proceedings of the SIGCHI Conference on Human Factors in*

Computing Systems (pp. 677–686). New York, NY, USA: ACM. https://doi.org/10.1145/1240624.1240730

Ishimaru, S., Kunze, K., Kise, K., Weppner, J., Dengel, A., Lukowicz, P., & Bulling, A. (2014). In the Blink of an Eye: Combining Head Motion and Eye Blink Frequency for Activity Recognition with Google Glass. In *Proceedings of the 5th Augmented Human International Conference* (p. 15:1–15:4). New York, NY, USA: ACM. https://doi.org/10.1145/2582051.2582066

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, *396*(6708), 228. https://doi.org/10.1038/24300

Kane, S. K., Wobbrock, J. O., & Ladner, R. E. (2011). Usable Gestures for Blind People: Understanding Preference and Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 413–422). New York, NY, USA: ACM. https://doi.org/10.1145/1978942.1979001

Kaptelinin, V., & Nardi, B. A. (2009). *Acting with technology: Activity theory and interaction design*. The MIT Press.

Karam, M., Lee, J. C., Rose, T., Quek, F., & McCrickard, S. (2009). Comparing Gesture and Touch for Notification System Interactions. In *Second International Conferences on Advances in Computer-Human Interactions, 2009. ACHI '09* (pp. 7–12). https://doi.org/10.1109/ACHI.2009.65

Karapanos, E., Jain, J., & Hassenzahl, M. (2012). Theories, methods and case studies of longitudinal HCI research. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (pp. 2727–2730). New York, NY, USA: ACM. https://doi.org/10.1145/2212776.2212706

Kato, H., & Billinghurst, M. (1999). Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In *2nd IEEE and ACM International Workshop on Augmented Reality, 1999. (IWAR '99) Proceedings* (pp. 85–94). https://doi.org/10.1109/IWAR.1999.803809

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press.

Kensing, F., & Blomberg, J. (1998). Participatory Design: Issues and Concerns. *Computer*

Supported Cooperative Work (CSCW), *7*(3–4), 167–185. https://doi.org/10.1023/A:1008689307411

Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B., Pagulayan, R. J., & Wixon, D. (2008). Tracking Real-time User Experience (TRUE): A Comprehensive Instrumentation Solution for Complex Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 443–452). New York, NY, USA: ACM. https://doi.org/10.1145/1357054.1357126

Kjeldskov, J., & Graham, C. (2003). A Review of Mobile HCI Research Methods. In L. Chittaro (Ed.), *Human-Computer Interaction with Mobile Devices and Services* (pp. 317–335). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-45233-1_23

Kohan, A., & O'Connor, B. P. (2002). Police Officer Job Satisfaction in Relation to Mood, Well-Being, and Alcohol Consumption. *The Journal of Psychology*, *136*(3), 307. https://doi.org/10.1080/00223980209604158

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., & Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. *Interacting with Computers*, *23*(5), 473–483. https://doi.org/10.1016/j.intcom.2011.06.005

Lanzara, G. F. (1983). *The Design Process: Frames, Metaphors, and Games*.

Law, E. L. C., & van Schaik, P. (2010). Modelling user experience - An agenda for research and practice. *Interacting with Computers*, *22*, 313–322. https://doi.org/10.1016/j.intcom.2010.04.006

Lieberson, S. (1991). Small N's and big conclusions: an examination of the reasoning in comparative studies based on a small number of cases. *Social Forces*, *70*(2), 307–320.

Lieberson, S. (1994). More on the uneasy case for using Mill-type methods in small-N comparative studies. *Social Forces*, *72*(4), 1225–1237.

Liu, J., Zhong, L., Wickramasuriya, J., & Vasudevan, V. (2009). uWave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing*, *5*(6), 657–675. https://doi.org/10.1016/j.pmcj.2009.07.007

Martin, B., Hanington, B., & Hanington, B. M. (2012). *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*.

Rockport Publishers. Retrieved from http://proquest.safaribooksonline.com.ezproxy.lib.vt.edu/9781592537563

Mayhew, D. J. (1998). The Usability Engineering Lifecycle. In *CHI 98 Cconference Summary on Human Factors in Computing Systems* (pp. 127–128). New York, NY, USA: ACM. https://doi.org/10.1145/286498.286575

Mayhew, D. J. (1999). *The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design*. Morgan Kaufmann.

McCrickard, D. S., Catrambone, R., Chewar, C. ., & Stasko, J. T. (2003). Establishing tradeoffs that leverage attention for utility: empirically evaluating information display in notification systems. *International Journal of Human-Computer Studies*, *58*(5), 547–582. https://doi.org/10.1016/S1071-5819(03)00022-3

McFarlane, D. C. (2002). Comparison of Four Primary Methods for Coordinating the Interruption of People in Human-Computer Interaction. *Human–Computer Interaction*, *17*(1), 63–139. https://doi.org/10.1207/S15327051HCI1701_2

McGrenere, J., Baecker, R. M., & Booth, K. S. (2002). An evaluation of a multiple interface design solution for bloated software (pp. 164–170). ACM.

McMahan, A. (2003). Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. *The Video Game Theory Reader*, *67*, 67–86.

McMahan, R. P., Bowman, D. A., Zielinski, D. J., & Brady, R. B. (2012). Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics*, *18*(4), 626–633.

McMahan, R. P., Gorton, D., Gresock, J., McConnell, W., & Bowman, D. A. (2006). Separating the effects of level of immersion and 3D interaction techniques. In *Proceedings of the ACM symposium on Virtual reality software and technology* (pp. 108–111). ACM.

Miles, M. B. (1979). Qualitative Data as an Attractive Nuisance: The Problem of Analysis. *Administrative Science Quarterly*, *24*(4), 590–601.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321–1329.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies* (Vol. 2351, pp. 282–293). International Society for Optics and Photonics.

Mitra, S., & Acharya, T. (2007). Gesture Recognition: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *37*(3), 311–324. https://doi.org/10.1109/TSMCC.2007.893280

Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., Wobbrock, J. O., & others. (2014). Reducing legacy bias in gesture elicitation studies. *Interactions*, *21*(3), 40–45.

Morris, M. R., Wobbrock, J. O., & Wilson, A. D. (2010). Understanding users' preferences for surface gestures. In *Proceedings of graphics interface 2010* (pp. 261–268). Canadian Information Processing Society. Retrieved from http://dl.acm.org/citation.cfm?id=1839260

Muller, M. J. (2009). Participatory Design: The Third Space in HCI. In *Human-Computer Interaction. Development Process* (Second, pp. 165–185). CRC Press.

Myers, G. J., Sandler, C., & Badgett, T. (2011). *The Art of Software Testing, 3rd Edition* (3rd ed.). John Wiley & Sons. Retrieved from http://proquest.safaribooksonline.com.ezproxy.lib.vt.edu/9781118133156

NFRO. (2014). National First Responders Organization | Who Is A First Responder? Retrieved August 24, 2017, from http://www.nfro.org/who.html

Nielsen, J. (1992). The usability engineering life cycle. *Computer*, *25*(3), 12–22. https://doi.org/10.1109/2.121503

Nielsen, Jakob. (2008). Usability ROI Declining, But Still Strong (Jakob Nielsen's Alertbox) [Web Page]. Retrieved January 1, 2006, from http://www.useit.com/alertbox/roi.html

Online Etymology Dictionary. (2016, March 11). Retrieved March 11, 2016, from http://www.etymonline.com/index.php?term=gesture&allowed_in_frame=0

Peattie, L. (2001). Theorizing Planning: Some Comments on Flyvbjerg's Rationality and Power. *International Planning Studies*, *6*(3), 257–262. https://doi.org/10.1080/713672898

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77.

Rabbi, I., & Ullah, S. (2013). A survey on augmented reality challenges and tracking. *Acta Graphica Znanstveni Časopis Za Tiskarstvo I Grafičke Komunikacije*, *24*(1–2), 29–46.

Rudd, J., Stern, K., & Isensee, S. (1996). Low vs. High-fidelity Prototyping Debate. *Interactions*, *3*(1), 76–85. https://doi.org/10.1145/223500.223514

Ryan, C., & Potts, L. (2015). Leading participant-centered research: an argument for taking a more strategic role as user experience architects. In *Proceedings of the 33rd Annual International Conference on the Design of Communication* (p. 20). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2775476

Salary: Entry Level Programmer. (2017a). Retrieved December 12, 2017, from https://www.glassdoor.com/Salaries/entry-level-programmer-salary-SRCH_KO0,22.htm

Salary: Researcher. (2017b). Retrieved December 12, 2017, from https://www.glassdoor.com/Salaries/researcher-salary-SRCH_KO0,10.htm

Sanna, A., & Manuri, F. (2016). A Survey on Applications of Augmented Reality. *Advances in Computer Science : An International Journal*, *5*(1), 18–27.

Schuler, D., & Namioka, A. (1993). *Participatory Design: Principles and Practices*. CRC Press.

Smith, S. D., & Caruso, J. B. (2010). *The ECAR study of undergraduate students and information technology, 2010*. Educause Colorado. Retrieved from http://net.educause.edu/ir/library/pdf/ekf/ekf1006.pdf

Soo Kyun Kim, Shin-Jin Kang, Yoo-Joo Choi, Min-Hyung Choi, & Min Hong. (2017). Augmented-Reality Survey: from Concept to Application. *KSII Transactions on Internet & Information Systems*, *11*(2), 982–1004. https://doi.org/10.3837/tiis.2017.02.019

Stake, R. E. (1995). *The Art of Case Study Research*. Thousand Oaks: Sage Publications.

Stake, R. E. (2009). The Case Study Method in Social Inquiry. In R. Gomm, M. Hammersley, & P. Foster, *Case Study Method* (pp. 18–26). 1 Oliver’s Yard, 55 City Road, London England EC1Y 1SP United Kingdom: SAGE Publications Ltd. Retrieved from http://methods.sagepub.com/book/case-study-method/d5.xml

Starman, A. B. (2013). The case study as a type of qualitative research. *Journal of Contemporary Educational Studies/Sodobna Pedagogika*, *64*(1), 28–43.

Streefkerk, J. W. (2011). Doing the right task: Context-aware notification for mobile police teams.

Retrieved from https://repository.tudelft.nl/islandora/object/uuid:c342f633-dfad-4ad7-bfe3-9dfe213fa7de/?collection=research

Streefkerk, Jan Willem, Wiering, C., van Esch-Bussemakers, M., & Neerincx, M. (2008). Effects of Presentation Modality on Team Awareness and Choice Accuracy in a Simulated Police Team Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52*(4), 378–382. https://doi.org/10.1177/154193120805200438

Tennant, G. (2001). *Six Sigma: SPC and TQM in Manufacturing and Services*. Gower Publishing, Ltd.

Thomas, G. (2011). A Typology for the Case Study in Social Science Following a Review of Definition, Discourse, and Structure. *Qualitative Inquiry*, *17*(6), 511–521. https://doi.org/10.1177/1077800411409884

Tucker, A. B. (2004). *Computer Science Handbook*. CRC press.

Tullis, T., & Albert, W. (2010). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics* (1st ed.). Morgan Kaufmann. Retrieved from http://www.amazon.com/Measuring-User-Experience-Technologies-ebook/dp/B005NZ7GS6/ref=tmm_kin_title_0

Van Krevelen, D. W. F., & Poelman, R. (2010). A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, *9*(2), 1.

Van Schaik, P., & Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers*, *23*(1), 18–32.

Vatavu, R.-D., Anthony, L., & Wobbrock, J. O. (2012). Gestures as point clouds: a $ P recognizer for user interface prototypes. In *Proceedings of the 14th ACM international conference on Multimodal interaction* (pp. 273–280). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2388732

Vickers, A. J., & Altman, D. G. (2013). Statistics Notes: Missing outcomes in randomised trials. *BMJ*, *346*, f3438. https://doi.org/10.1136/bmj.f3438

W. Gary Howard, Heather Howard Donofrio, & James S. Boles. (2004). Inter-domain work-family, family-work conflict and police work satisfaction. *Policing: An International Journal*, *27*(3), 380–395. https://doi.org/10.1108/13639510410553121

Walker, M., Takayama, L., & Landay, J. A. (2002). High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes when Testing Web Prototypes. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(5), 661–665. https://doi.org/10.1177/154193120204600513

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1083–1092). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1518866

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*(Supplement C), 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

Yin, R. K. (1981). The Case Study Crisis: Some Answers. *Administrative Science Quarterly*, *26*(1), 58–65.

Yin, R. K. (1984). *Case study research: design and methods*. Sage Publications.

Yin, R. K. (1994). Discovering the future of the case study method in evaluation research. *Evaluation Practice*, *15*(3), 283–290. https://doi.org/10.1016/0886-1633(94)90023-X

Zhou, F., Duh, H. B.-L., & Billinghurst, M. (2008). Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* (pp. 193–202). Washington, DC, USA: IEEE Computer Society. https://doi.org/10.1109/ISMAR.2008.4637362

8: Appendix A

Storyboard 2.0 Figure

9: Appendix B

Control Interface Figure