

# A Systems Approach for Augmented Reality Design

Andrea K. Webb<sup>(✉)</sup>, Emily C. Vincent, Pooja Patnaik, and Jana L. Schwartz

Draper, Cambridge, MA, USA

{awebb, evincent, ppatnaik, jana}@draper.com

**Abstract.** Effective ways of presenting digital data are needed to augment a user's experience in the real world without distracting or overloading them. We propose a system of systems approach for the design, development, and evaluation of information presentation devices, particularly augmented reality devices. We developed an evaluation environment that enables the synchronized presentation of multimodal stimuli and collection of user responses in an immersive environment. We leveraged visual, audio, thermal, and tactile information presentation modalities during a navigation and threat identification task. Twelve participants completed the task while response time and accuracy data were collected. Results indicated variability among devices and pairs of devices, and suggested that information presented by some pairs of devices was more effective and easily acted upon than that presented by others. The results of this work provided important guidance regarding future design decisions and suggest the utility of our system of systems approach. Implications and future directions are discussed.

**Keywords:** Augmented reality · Situation awareness · System of systems · Immersive environment

## 1 Introduction

We are surrounded by digital data and are faced with challenges in how to best acknowledge and use that data. Too much data or conflicting/confusing data has far-reaching implications and may lead to events such as friendly fire incidents, clinical errors (e.g., operating on the wrong body part or side), or civilian accidents (e.g., car accidents because of cell phone distractions). What is needed is a way to effectively present digital information that augments what the user experiences in the real world without distracting or overloading the user or making their task more difficult.

In order to address these types of challenges, a multidisciplinary and multifaceted approach is needed. The domain of Human Systems Technology requires a spectrum of user modeling, observation, and measurement techniques. We select techniques spanning from expert assessment through rigorous human subjects testing in order to meet the needs of a particular design challenge. Both ends of this spectrum have strengths and weaknesses: expert judgment is an interpolation based on experience and is thus limited in scope; human subjects testing – particularly in the applied research space – has come under scrutiny due to difficulties in reproducing results [1]. Taken together, these challenges point to the need for a different approach to leverage expert or user response to guide design decisions.

There are a number of systems engineering approaches that can help guide design decisions. There are several systems analysis tools and techniques that can be applied across the spectrum of design maturity [2]. However, both modeling and measurement are less mature for human-in-the-loop systems. Current laboratory measures and methods largely remove technology from its intended environment. This is problematic because out of context performance data tends to provide minimal design insight. Approaches we have leveraged to include and assess context include testing a single design across a number of user cases [3]; this approach provides robust findings but can be expensive and time consuming. Another approach is to measure multiple channels of data continuously during completion of a task, rather than relying solely on pre/post assessments [4]. An additional challenge is that we wish to provide feedback at the speed of design (e.g., over days), as opposed to the speed of academic publication (e.g., months or years).

How then do we develop an evaluation environment to support the design and development of a system of novel wearables, with the purpose of providing the user with a sense of immersive situation awareness (SA)? Our initial operational requirements were to measure situation awareness as a function of presentation and context. We define SA to be measured by accuracy moderated by response time, and our goal is to develop designs to maximize SA.

In order to ground our research in a specific use case with real-world implications, we selected a ground infiltration mission as our primary concept of operations. These users have several tasks they are trying to do concurrently while navigating their mission (e.g., watching for hazards, communicating with other team members). They have a number of tools available to them, such as handheld devices for navigation and monitoring, but they are often challenged to maintain a balance of heads down and heads up time to successfully execute their mission.

Our broader interest is to provide an evaluation solution that is useful and cost effective for the crowded small business and start-up space in augmented reality, university researchers, and the simulation and test community. This paper details the first implementation of our approach and provides important guidance for future revisions.

## 2 Method

### 2.1 Participants

Participants were 12 Draper employees. Of these, 8 (67 %) were male and 4 (33 %) were female. Participant age ranged between 23 and 52 ( $M = 34.33$ ,  $SD = 9.38$ ). Due to data collection issues, not all participants experienced all stimulus conditions. As such, statistical analysis was not performed; summary measures are reported below.

### 2.2 Testing Environment

We have developed a lab for the rapid design and evaluation of directed attention and performance augmentation devices that can be tested against metrics relevant in the operational space (accuracy, response time, psychophysiological response). The

physical environment is comprised of eight identical ambient presentation panels, each of which can present local audio and visual stimuli. The baseline configuration for the panels is a uniform closed shape (an octagon), but the layout is dynamically configurable to provide a variety of orientations. This modular design provides flexibility for testing numerous devices and configurations with tasks of various degrees of flexibility and noise. Its scalability (from 1:8 or more panels) allows for enhancement with minimal added investment, using the same presentation and data collection infrastructure.

The physical octagon environment is grounded in a lab-centered, lab-fixed coordinate system. All lab infrastructure and all hardware units under test are referenced to this 3D coordinate space. Communication among lab and test devices, including timing synchronization, high level commands, and interdevice messaging, is enabled by the Draper Network Architecture for Java software framework (DNA4J). DNA4J is a pure Java solution enabling development on any platform with support for Java 1.6 or later. It has been tested on Windows, Linux, OSX, and also supports the Raspberry PI Raspian OS.

In addition to uniform presentation of lab and digital stimuli, the octagon further enables concurrent recording of user response. Each panel is equipped with a motion capture camera that can capture the 3D position of multiple points in a scene. To date we have collected rigid body data on the user's head, torso, and handheld devices. The DNA4J framework further enables time synchronized collection of task and user responses.

### 2.3 Test Devices

We performed an integrated data collection that leveraged four different input devices: augmented reality glasses that included a heads-up display for presentation of visual information, a wrist-worn thermal device for presentation of warm and cool temperatures, standard earbuds for presentation of audio information, and a chest-worn torque tactile device (TTD) that used torque for presentation of information. We also leveraged a handheld Android phone for presentation of visual information. Participants wore each device singly (Android handheld, AR glasses, thermal, audio, TTD) and in pairs (e.g., AR + TTD, thermal + audio, AR + Android handheld) for the trials discussed below.

### 2.4 Test Protocol

Participants were asked to complete a navigation and threat identification task in the octagon environment (see Fig. 1). Participants were presented with black and yellow circles on the handheld Android device. Participants were instructed to act upon the yellow circles by pressing the target button on the handheld device (see Fig. 2). After some trials, a threat was signaled on one of the octagon panels (northwest, southwest, southeast, northeast). There was no visual threat actually presented on the octagon panel. Participants were presented with a cue on each device or pair of devices signaling them to which panel to turn to then acknowledge the threat. For example, if the threat was on the southeast panel, the AR glasses would display an arrow pointing in the southeast

direction, and the handheld Android device would present a symbol on the southeast panel on the display. Similarly, if thermal and TTD were paired, the thermal device would present a warm temperature, and the TTD would send signals to nudge the participant to the southeast panel. Participants completed 4 trials per device and device pairing (44 trials total) in the context condition in which a threat always appeared after a yellow circle, and 44 trials in the no context condition in which a threat could appear after a yellow or black circle. Response time and accuracy was assessed for each trial.



Fig. 1. The octagon environment

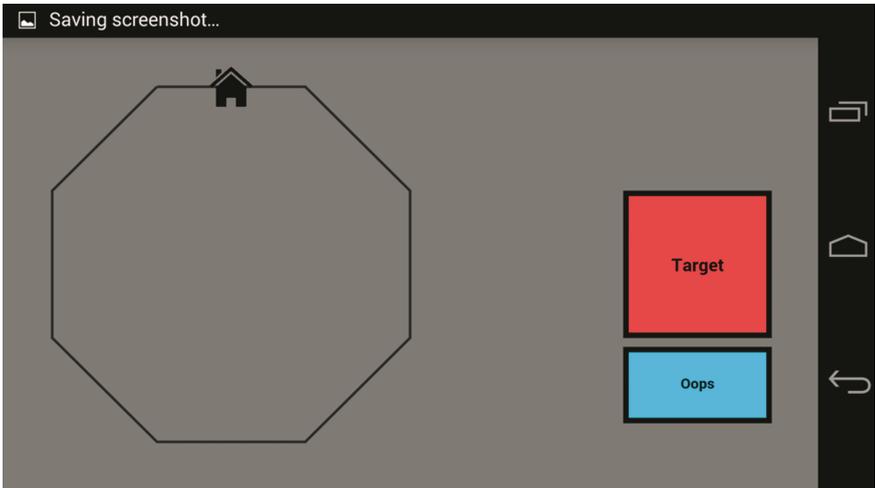
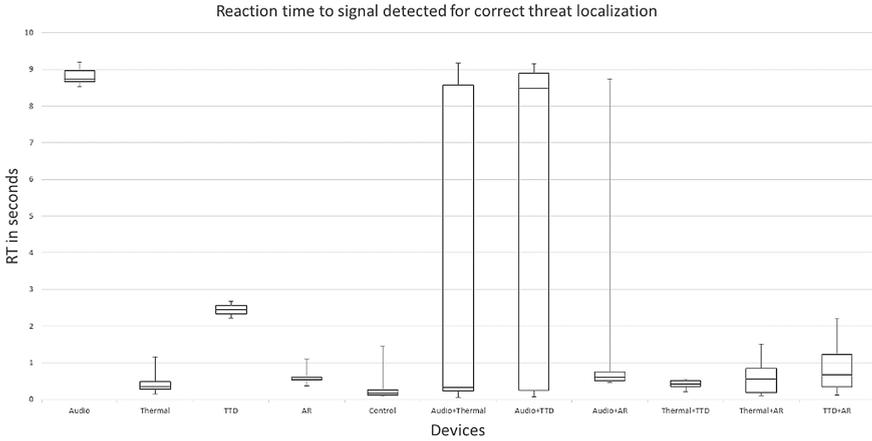


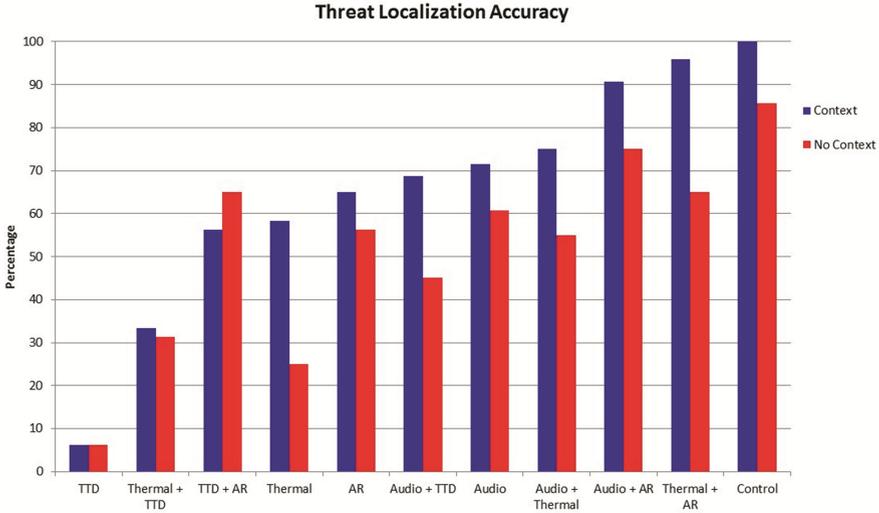
Fig. 2. Android handheld device interface for the navigation and threat identification task

### 3 Results

Response time and accuracy were assessed for each of the devices and pairs of devices. Response time data for trials where the threat was correctly localized are presented in Fig. 3, and accuracy data are presented in Fig. 4.



**Fig. 3.** Reaction time results for each device and pairs of devices for correct trials



**Fig. 4.** Threat localization accuracy for each device and pairs of devices

It can be seen in these plots that there was a large amount of variability across devices – it was much easier to correctly detect a threat with cues from some of the devices (i.e., Android handheld and AR glasses) and pairs of devices (i.e., AR

glasses + thermal, audio + AR) than it was for other devices (i.e., TTD). It also can be seen that response times were much longer for some devices (i.e., audio) than for others (i.e., AR). An examination of participant self-report responses and a close examination of the cues indicated that the audio cues took longer to complete than did the cues from other devices, and participants would wait until completion of the presentation before indicating the threat location. Participant self-report data also revealed that threat localization cues were easiest to detect with the AR glasses and Android handheld, whereas cues were most difficult to detect with the thermal, TTD, and audio devices. These self-report results are not unexpected and confirm the need to identify intuitive cues or provide learning trials (i.e., if using thermal cues, participants would need time to learn that cold = left, and hot = right). Despite the differences in response time and accuracy across devices and pairs of devices, the results were important for guiding future design decisions, as discussed below.

## 4 Discussion

Through our current work, we have come to believe that a true augmented reality system will encompass more than a set of AR goggles. In order to augment a human in a way that will be intuitive and actionable, we must move beyond focus on the visual and include multiple sensory modalities. This implies a focus on design of systems of systems that must work in synchronization across a user's senses. In order to achieve this tight synchronization we have found that applying a systems engineering lens enables us to better examine the interactions and dependencies between different display systems.

Our systems engineering approach led us to start building our test environment at the beginning of the project, co-incident with designing our system of displays. Building a test environment from scratch had a number of benefits for our team. We had full control of the system, which enabled us to make it ideal for the scenarios we were testing. It was large enough to block out the rest of the room and allow the participant to move around in which, when combined with 360 degrees of projection surface, provided enough of an immersive environment for the participant to accept it as "reality." Creating an acceptable "reality" against which to overlay augmentation is crucial for any augmented reality testing endeavor.

We did find it was a significant effort to build a 360 degree immersive environment from scratch. While the octagon that we built meets our needs and was on the whole less expensive than buying a CAVE system, there are some downsides to its design. It is less clean/professional looking and it took more time and labor on our end. However, the Octagon has the added bonus of flexible configuration. As noted previously, it does not need to remain in an octagon formation, and our future plans involve using those panels in different configurations.

We also found that while our systems approach recognized extensibility and a "system of systems" concept, our software implementation of the messaging between systems had to undergo multiple iterations to meet this goal. Our initial implementation had each device implementing its own messaging scheme. This introduced a significant

burden to introducing new devices, which was one of the primary goals of this work. As such, we iterated on this software infrastructure in order to reduce the barrier to entry to an acceptably small level of effort that did not require expert support.

Through our desire to truly compare each device on a level playing field, we designed a single information scenario that all devices needed to display. This caused our results to show that some devices had better performance than others. However, we believe that this was due to not designing for each devices' specific strengths. We have begun to explore the differences in information type and style – such as how suddenly the information is available and how quickly the user needs to be aware of and react to that data. The goal for a future experiment phase is to create a holistic system design where each device plays to its strengths, which we recommend as an approach to AR systems.

## 5 Path Forward

There are a number of goals for our future work in this area. Our high level goal is to identify good combinations of systems that enable augmented reality in multiple fielded scenarios, ranging from the battlefield to the operating room to the analyst workstation. Our test environment approach needs to enable us to quickly integrate new display systems as they become available and to quickly configure and run tests on multiple combinations of new and existing systems. To this end, we plan to continue to build flexibility and extensibility into our software framework. This includes well documented APIs, generalizable messaging, and a general focus on adding extensibility to the software architecture. One specific action we are undertaking is to ensure that our messaging infrastructure sends information or data rather than commands on how to display that information. This will enable each device to make decisions on how best to display information and will de-centralize the intelligence of the system, making it easier to integrate new components without requiring changes in the core software.

We also have considered increasing the immersiveness of the testing environment, in order to increase the user acceptance of the “reality” with which our devices are augmented. Additional study is needed to determine how the immersion of our environment impacts the experimental results and generalizability of findings to the real world. Some augments to the immersiveness that we've considered are: full 360 degree motion environment though a gaming engine such as Unity, adding a 360 degree treadmill to enable participant movement, ambient sound, and ambient thermal. Future work will explore how close to a fully enclosed CAVE the environment needs to be or if a minimalist approach produces the same results.

Another approach to increasing the immersiveness of the test environment is to open up the octagon. The modularity of the octagon enables multiple configurations – our next planned experiment involves using a subset of the 8 panels to create a 180 degree environment that will simulate a driving scenario. We also have planned to break individual panels out and place them in an actual physical environment – such as a house or obstacle course – so that we can combine physical features and rooms with virtualized panels, similar to a haunted house experience.

Perhaps most importantly, we plan to bring new devices into our system of augmented reality devices and test environment. We are exploring multiple haptic feedback systems, such as vibrotactile boots, belts, and a headband. We also have started acquiring multiple brands of AR goggles and smart glasses. We continue to explore other sensory modalities that could be applicable, such as smell and taste. Lastly, we want to expand our testing to include immersive user input devices – devices that enable the user to communicate back to the system of systems in a manner that does not distract them from or interrupt their physical world task. These may initially include gesture, through use of hardware such as data gloves or vision, physical controls such as buttons, knobs, and touchscreens, implicit interactions such as gaze or head movement, and voice control, both audible and sub-vocal. We believe that immersing the user in physical and digital information is not enough to keep them from being heads down in their technology. We must also free them from the burden of inputting data into their augmentation systems. Taken together, these future research plans will provide a rich set of data and design guidelines that can be used by future developers of augmented reality devices and systems.

## References

1. Collins, F.S., Tabak, L.A.: NIH Plans to enhance reproducibility. *Nature* **505**, 612–613 (2014)
2. Borer, N.K., Odegard, R.G., Schwartz, J.L., Arruda, J.R.: A unified framework for capturing concept development methods. In: *Proceedings of the IEEE Aerospace Conference, Big Sky, Montana* (2009)
3. Martin, D.J., Martin, J.Z., Webb, A.K., Horgan, A.J., Marchak, F.M.: *Malintent Detection in Primary Screening Utilizing Noncontact Sensor Technology* (2016, Manuscript in preparation)
4. Poore, J.C., Webb, A.K., Cunha, M.G., Mariano, L.J., Chappell, D.T., Coskren, M.R., Schwartz, J.L.: Operationalizing engagement as coherence to context: a cohesive physiological measurement approach for human computer interaction. *IEEE T. Affect. Comput.* (in press)