Investigating a Virtual Reality-based Emergency Response Scenario and Intelligent User Interface for First Responders

Randall Spain[1], Jason Saville[1], Barry Liu[1], Donia Slack[2], Edward Hill[2], John Holloway[2]
Sarah Norsworthy[2], Bradford Mott[1], & James Lester[1]
[1]North Carolina State University, Center for Educational Informatics, Department of Computer Science
[2]RTI International, Center for Forensic Sciences

Because advances in broadband capabilities will soon allow first responders to access and use many forms of data when responding to emergencies, it is becoming critically important to design heads-up displays to present first responders with information in a manner that does not induce extraneous mental workload or cause undue interaction errors. Virtual reality offers a unique medium for envisioning and testing user interface concepts in a realistic and controlled environment. In this paper, we describe a virtual reality-based emergency response scenario that was designed to support user experience research for evaluating the efficacy of intelligent user interfaces for firefighters. We describe the results of a usability test that captured firefighters' feedback and reactions to the VR scenario and the prototype intelligent user interface that presented them with task critical information through the VR headset. The paper concludes with lessons learned from our development process and a discussion of plans for future research.

INTRODUCTION

Virtual environments (VE) and other forms of digitally mediated experiences, including virtual reality (VR), simulations, and serious-games, have been used for years to train personnel in domains spanninng defense (Taylor & Barnett, 2013), aviation (Oberhauser, Dreyer, Braunstingl, & Koglbauer, 2018), medicine (Aïm, Lonjon, Hannouche, & Nizard, 2016), aerospace (Little, 2017), and homeland security (Griffith, Ablanedo, & Dwyer, 2017). These platforms offer a number of affordances to support safe and effective training. In addition to being an immersive training medium, VR has emerged as an important tool in product development and user experience research. By leveraging high fidelity graphics, engaging scenarios, and expressive displays, researchers can use VR to envision and systematically investigate alternative user interface designs in safe, replicable, and controlled environments.

A distinctive application of VR for supporting user experience research is its use as a platform for investigating how prototype displays should present information to end-users. Heads-Up Displays (HUDs) offer a promising medium for providing end-users with task-critical data, particularly first responders. For instance, firefighters could receive information about air tank levels, thermal readings, and navigation all at a glance through an interactive HUD. However, as data becomes increasingly available, a critical challenge facing HUD designers is ensuring firefighters receive the right information, in the right format, at the right time through without imposing significant levels of mental workload or frustration. Moreover, firefighters must be able to use interact with data presented through these displays efficiently and effectively without committing unnecessary errors due to poor design.

In this paper, we describe the development of a VR-based emergency response scenario that serves as a testbed for evaluating how intelligent user interfaces (IUI) should be designed to support firefighters. IUIs leverage artificial intelligence techniques such as machine learning, natural language processing, data mining, knowledge representation, and reasoning to provide end-users with tailored support and information. Our interests are centered on investigating whether IUIs that offer speech interaction capabilities and adaptive informatics can enhance task performance for firefighters and facilitate more natural human-computer interaction. The paper also describes the results of a usability evaluation that captured firefighters' feedback on the VR scenario and a prototype IUI. We conclude with lessons learned from the development process and plans for future research

THE VIRTUAL REALITY SCENARIO

Over the past three years, the Public Safety Communications Research (PSCR) Division of the National Institute of Standards and Technology (NIST) has been investigating how VR and augmented reality (AR) can be used to evaluate user interfaces for the public safety community. In line with this goal, our team has developed a VR-based scenario that serve as an environment for evaluating how IUI principles and AR display configurations should be designed to better assist firefighters when responding to emergencies.

The VR environment was developed in Unity 3D and simulates a fire scenarios similar to the one described in the "Incident Scenarios Collection for Public Safety Communications Research" report (Choong, Dawkins, Greene, & Theofanos, 2017) in which first responders conducted a coordinated multi-unit search and rescue mission for a disabled rail car in a metro tunnel while facing heavy smoke and sporadic radio communication.

The team engaged in several key activities to design and develop the VR metro environment and supporting scenarios. These activities closely aligned with an Analyze, Design, Develop, Implement and Evaluate (ADDIE) instructional design model (Branch, 2009) wherein the team began by reviewing publicly available reports of metro incidents from the National Transportation Safety Board (NTSB), metro incident standard operating procedures (SOPs), and emergency response guides to identify the tasks and procedures associated with responding to metro incidents and to better understand the operational environment. The team also visited the Washington Metro Emergency Response Training Facility and interviewed firefighters and training mangers to discuss SOPs and to identify critical elements to include in the VR metro scenarios. The project team used this information, along with input gathered from firefighters on a baseline model of the VR environment, to construct a rich VR-based subway fire scenario to serve as the IUI testbed for our research (Figure 1).

The current iteration of the VR-based metro emergency response scenario includes three missions. In each mission, users take on the role of a firefighter whose objective is to navigate through the metro station, find the disabled metro train, and assist any passengers that need rescuing. Users are instructed to complete seven tasks as they work towards completing the mission's goals. These include: (1) opening the metro tunnel gate via a barrel key so that passengers can egress down the catwalk; (2) confirming with a non-player character that the third rail is de-energized; (3) de-energizing the third rail at emergency box, if needed; (4) locating the metro car; (5) opening the metro train door using the barrel key; (6) entering the train and extracting any unconscious passengers; and (7) evacuating with the passenger to the mezzanine as quickly as possible. Scenario configurations such as the starting location and placement of the missing metro train and injured passenger can be altered for each mission. The density of smoke in the subway station and the amount of time users have to complete the mission can also be manipulated using a configurations toolbox.

Figure 1. VR scenario metro environment.

The VR scenario also includes a secondary radio monitoring task that requires users to listen to a simulated radio channel in the VR headset and respond by pressing the trigger on the left controller whenever a target stimulus is presented. The secondary task is intended to be performed concurrently with the primary task objectives outlined in each scenario. The auditory stimuli used for the task simulates a dispatcher issuing call signs to response units. The user's task is to respond whenever he or she hears a dispatch call in which the company and battalion number total to an even number (e.g., "Battalion 1, 3rd Truck Company"). In the scenario, new radio dispatch calls are presented every 15 seconds and users have a 12 second window to perform the mental calculation and make a response decision. The number of "target signals" for each session as well as the interstimulus interval can be manipulated by the researcher using the toolbox configuration. Users receive points for making a correct response (responding when a target signal is present or not responding to a non-target signal) and lose points when they make an incorrect response (i.e., false alarm or miss). Accuracy and response time data are recorded for each trial and logged to a database. At the conclusion of each mission, users receive feedback on how many objectives they completed, their performance accuracy on the radio monitoring task, and how efficiently they completed the scenario.

In addition to the three scenarios, the VR module includes a controller tutorial that provides users with direct instruction on how to use the VR controllers to navigate in the environment, manipulate and pick up objects, and activate the prototype IUI which provides users with task-critical data through the VR headset's HUD.

VR controller options. The VR scenario includes two options for navigating the subway environment. The first involves point-and-click teleportation wherein the user aims the controller in a specific direction and presses the VR control pad to teleport to the aimed location. The second is motion based and requires participants to depress the trackpad on the right VR controller and swing their arms forward and backward in a walking motion. Users control their locomotion rate by swinging their arms faster or slower. For the evaluation described in this paper, participants used the walk-in-place motion to navigate in the environment. Users also use the VR controller to grab and manipulate objects in the environment. To climb out of the railway bed, users perform a hoisting gesture wherein they reach forward to the platform ledge, press and hold the right trigger, and then move their arms downward as if climbing up onto the ledge.

HUD controls. The scenario also includes a prototype IUI that provides users with task-relevant informatics through the VR headset. The informatics include: navigational information, a dynamic gauge that provides an estimate of how much time is remaining until the air tank is depleted, a dynamic temperature gauge that shows environmental temperature and uses principles of urgency mapping (Baldwin & Lewis, 2014) to display critical changes in temperature readings based on the user's proximity to a fire, an "edge detection" feature that facilitates the visualization of object outlines in zero-visibility conditions, text translations of radio communication messages, and prompts that remind participants of tasks they need to complete.

The navigation display and air tank and temperature gauges are presented to users in a fixed location in the HUD and no user action is required to access this information (Figure 2). Users summon the radio communication information and mission prompts by depressing and holding the left grip button on the VR controller. Users can activate the edge detection feature by depressing the left control pad. Users can also use voice commands to summon or dismiss informatics presented through the HUD.

Figure 2. Example of informatics presented through the HUD.

USABILITY EVALUATION

To test the feasibility of the VR scenario for facilitating IUI research, the project team conducted a usability evaluation with subject matter experts. This evaluation aided in the development of the VR build by allowing firefighters to experience the scenario and provide feedback on the environment, HUD, and controller tutorial. The research team also gathered feedback on the utility of the HUD for completing the simulated missions. For this exercise, the speech interface was disabled so users summoned and dismissed informatics presented through the HUD using VR controllers as described previously.

Participants

A total of 8 firefighters from the Washington, DC metro area participated in the usability testing session and each firefighter completed at least one mission from the VR emergency response scenario. Participants were recruited from a training session that took place at the Washington Metro Emergency Training Facility. Participants self-reported having low levels of experience with VR and video games.

Apparatus

Participants completed the session in a training room that included three VR workstations and tables for taking notes and administering online surveys (Figure 3). Each workstation contained a laptop with an Intel Core i7-8750K processor with 32 GB of RAM and a Nvidia RTX 2080 GPU and an HTC VIVE Pro headset with lighthouses and controllers.

Procedure

After providing consent to participate, firefighters were briefed on the purpose and goals of the session. Then, they reviewed the feedback forms that the research team asked them to complete. The feedback forms contained sample questions for firefighters to keep in mind while participating in the usability testing session. Questions were designed to capture feedback about any technical issues that participants experienced as well as to gather feedback on the clarity of the mission instructions, the usability of controllers, the fidelity of the scenario, the design of the HUD, and any issues that impacted their user experience. Firefighters were informed that a member of the research team would guide them through the exercise, observe their interactions with the VR scenario, and take notes and record any comments they had regarding the VR scenario. Firefighters were also informed that upon completing the testing session, they would be asked to complete several surveys to gather their final comments and thoughts about the VR scenario and IUI.

After receiving these instructions, participants donned the VR headset with the assistance of a research staff member and began the testing session. Each firefighter completed the controller tutorial prior to completing at least one of the three available search and rescue missions. Each mission required firefighters to complete the seven tasks described previously as they navigated through the metro station. Firefighters could access task-critical information, including navigation support, task reminders, and radio-to-text support, through the prototype intelligent HUD using the VR controllers. In addition to completing the primary task of rescuing any passengers that needed assistance, firefighters performed the secondary radio monitoring task described previously. Participants were instructed to complete each mission as quickly and accurately as possible and to provide feedback on any issues or problems they experienced during the mission.

Figure 3. Firefighters participating in the testing session.

After completing the VR mission, firefighters completed a brief demographic questionnaire, a graphical user interface (GUI) questionnaire, and an exercise questionnaire (Singer & Knerr, 2010). The GUI questionnaire contained 28 items that assessed reactions to the VR emergency response scenario and prototype IUI. Participants rated items such as, “Were the VR controller buttons and their associated functions easy to remember?”, “How realistic were the visual effects of the smoke in the scenario?”, and “How realistic were your capabilities in the VR environment for communicating with others?” using a Likert scale on which higher ratings indicated better usability and higher fidelity.

The exercise fidelity questionnaire contained nine items that addressed the training effectiveness potential of the scenario (Singer & Knerr, 2010). The research team also added six items to the questionnaire to assess the perceived utility of the HUD's informatics for completing the mission. Ratings were made on a 5-point Likert scale on which higher values indicated higher mission utility.

RESULTS

Graphical User Interface Questionnaire

A summary of firefighters' responses to the GUI questionnaire is presented in Table 1. Results were analyzed by calculating the item means for each dimension.

Table 1: *Mean Reviewer Ratings from the GUI Questionnaire*.

Dimension	Mean	SD	Min	Max
User Interface Satisfaction*	4.94	1.20	3.00	5.86
Scenario Fidelity	3.99	0.94	2.43	4.86
Manipulating Objects	3.50	1.17	1.50	4.50
Learnability and error recovery	3.95	0.19	3.80	4.20

*Ratings ranged from 1 to 7; all other ratings ranged from 1 to 5.

User interface satisfaction. Results showed that, on average, firefighters rated the VR scenario's user interface as being easy to understand and the VR controller buttons gestures and their associated functions as being easy to use. The respondents noted that it was fairly easy to learn the locomotion control system.

Scenario fidelity and realism. Firefighters reported that the metro station, metro tunnels, train cars, and smoke in the scenario were appropriately realistic for this type of simulation. When asked how realistic the VR application's capabilities were across different interaction functions, the majority of respondents stated moving in the environment was highly realistic, that communicating with others was somewhat realistic, and that performing physical tasks was moderately realistic. These ratings are encouraging and suggest the VR scenario contains an appropriate level of fidelity.

Manipulating objects and using gestures in the scenario. The majority of firefighters stated it was easy to manipulate objects in the environment and that using simple gestures to accomplish tasks in the scenario was appropriate for this type of exercise. The gesture for climbing (e.g., hoisting oneself from the railway to the platform) received negative reactions, though, with a majority of respondents stating the gesture seemed artificial and difficult to learn. Results suggest participants in our target sample who had low levels of experience using VR systems found it difficult to properly execute this gesture.

Learnability and error recovery. Overall, firefighters expressed that once they learned how to use the VR interface, they could easily focus on accomplishing the mission and performing the tasks required of the exercise.

Exercise Fidelity Questionnaire

The exercise fidelity questionnaire contained items that assessed whether the VR scenario evoked similar levels of coordination, stress, and fidelity as compared to field training exercises. Responses were made using a 5-point Likert scale, such that a response of 1 indicated "much lower," 3 indicated "about the same," and 5 indicated "much higher," as compared to field training. Items that examined the viability of the informatics presented through the HUD for completing the mission were made using a 5-point scale. We analyzed responses at the item level for both measures.

VR scenario compared to field training. When asked how the VR scenario compared to field training exercises with regard to realism and complexity, respondents rated the scenario as being lower in the time required to perform the exercise (M = 1.75) and much lower with regard to the amount of task induced stress (M = 1.50). Respondents also noted that the scenario required slightly lower levels of unit coordination (M = 2.25) and less radio communication (M = 2.00) compared to field training. These results are not unexpected given the scenario was not designed to induce high levels of task induced stress, high levels of unit coordination, or require participants to communicate with others. Importantly, firefighters in this sample rated the scenario as being only slightly lower than field training exercises in regard to the level of task complexity (M = 2.50) and mental demands required for completing the task (M = 2.50). These results suggest that the scenario's baseline tasks are inducing a moderate level of task and psychological fidelity.

HUD informatics. When asked how useful the informatics displayed through the HUD were for completing the mission, respondents noted that the navigation display was moderately effective for completing the mission (M = 3.84). Respondents also stated the radio communication display (M = 3.00) and data driven mission prompts (M = 2.80) were somewhat beneficial to their mission performance. The edge detection feature was only slightly effective for completing the mission. It is important to note that the scenario used in the testing session did not simulate zero-visibility conditions, thus making the utility of the edge detection feature less useful when compared to use in a smoke-filled environment. Overall, these reactions suggest that the informatics presented through the HUD offer the potential to benefit performance in the emergency response scenario.

Feedback Collected During User Trials

Despite gathering mainly positive reactions to the VR scenario and user interface from the questionnaires, the research team did identify several items and areas of the VR scenario and HUD that could be improved. A major observation during the testing session was that it took participants significantly longer to complete the controller tutorial than the team expected. Many participants reported having a low level of experience using VR systems, and for these individuals, it took up to 15 minutes to complete the VR controller tutorial and radio monitoring task tutorial (the project team anticipated it would take five minutes to complete the tutorial). Several participants noted that it was hard to read the instructional text and recall the VR controller buttons and gestures needed to complete the tutorial tasks. Others noted that it would be helpful to repeat each tutorial task several times prior to advancing to the next task. The project team has since addressed these issues by redesigning the tutorial to include character animations of the gestures and buttons that participants need to utilize to complete the actions in the tutorial. The instructions have also been segmented to reduce the amount of extraneous cognitive load placed on participants while learning the controller actions.

DISCUSSION AND LESSONS LEARNED

A primary goal of the usability evaluation was to test the feasibility of the VR scenario for facilitating IUI research. Despite only collecting data from a small sample of firefighters, the research team identified several strengths of the VR scenario and several areas for improvement. While caution should be exercised in interpreting results of a small sample study, this method may be preferable to the alternative of conducting a single evaluation session (Savage-Knepshield, 2007) where feedback is only gathered once from a larger sample. Our design approach continues to utilize an iterative process wherein feedback is repeatedly collected from firefighters to validate incremental design changes to the VR scenario.

One of the lessons learned from this effort is that gathering input from the target sample is incredibly valuable but also requires working within the constraints of firefighters' training schedules which means usability evaluations must take place quickly and efficiently. Moving forward, it will be important to consider this balance when collecting data from additional firefighters.

Our future scenario development plans include adding an alternative secondary task to the scenario that will require users to monitor radio chatter for a multi-unit coordinated response and respond to situation awareness prompts. This updated secondary task protocol will require users to keep information regarding the location of coordinating units in working memory longer than the current secondary task which serves as a simple mental arithmetic task. Further, additional input modalities are being added to the scenario that will allow users to use natural language and an AR-oriented system interface to activate and dismiss HUD informatics rather than using VR controller buttons. Including these additional input modalities will allow the research team to identify when and how often end-users summon HUD informatics and will allow the team to use data-driven approaches to develop models that can provide task-related informatics to end-users based on their inferred goals.

In conclusion, VR offers a promising medium for conducting user experience and interface research. It is important to note that the study described in this paper only evaluated the usability and feasibility of the VR scenario and HUD informatics; it did not directly address the relative effectiveness of a particular UI or the immersion of the scenario. Future research will focus on comparing input modalities (e.g., speech interface vs. VR controller) for interacting with the informatics available through the HUD to identify whether the affordances offered by IUIs improve system interactions and mission performance.

ACKNOWLEDGMENTS

The research described herein has been sponsored by the National Institute of Standards and Technology under grant number 60NANB18D148. The statements and opinions expressed in this article do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

Aïm, F., Lonjon, G., Hannouche, D., & Nizard, R. (2016). Effectiveness of virtual reality training in orthopedic surgery. *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, *32*(1), 224-232.

Baldwin, C. L., & Lewis, B. A. (2014). Perceived urgency mapping across modalities within a driving context. *Applied Ergonomics*, *45*(5), 1270-1277.

Branch, R. M. (2009). Instructional design: The ADDIE approach (Vol. 722). Springer Science & Business Media.

Choong, Y. Y., Dawkins, S., Greene, K., & Theofanos, M. (2017). Incident scenarios collection for public safety communications research. (NISTIR 8181). https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8181.pdf

Griffith, T., Ablanedo, J., & Dwyer, T. (2017). Leveraging a virtual environment to prepare for school shootings. In *International Conference on Virtual, Augmented and Mixed Reality* (pp. 325-338). Springer, Cham.

Little, W. (2017). The NASA Augmented/Virtual Reality Lab: The State of the Art at KSC.

Oberhauser, M., Dreyer, D., Braunstingl, R., & Koglbauer, I. (2018). What's real about virtual reality flight simulation? Comparing the fidelity of a virtual reality with a conventional flight simulation environment. *Aviation Psychology and Applied Human Factors*, *8*(1), 22-34.

Savage-Knepshield, P. A. (2007). Usability testing for rapid fielding with small N s: Lessons learned during an army operational field experiment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 51, No. 26, pp. 1622-1626). Sage CA: Los Angeles, CA: SAGE Publications.

Singer, M. J., Knerr, B. W. (2010). Evaluation of a game-based simulation during distributed exercises. (Research Report 1931). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences

Taylor, G. S., & Barnett, J. S. (2013). Evaluation of wearable simulation interface for military training. *Human factors*, *55*(3), 672-690.