Comparing Firefighters' Perceived Workload Using 2D vs. 3D Building Plans to Support Emergency Response Preplanning in a Simulated Fire Scenario

Katelynn A. Kapalo*
Christian Regenhard Center for Emergency Response Studies
Kevin Pfeil†
University of North Florida
Joseph Bonnell‡
Phoenix Fire Department
Joseph LaViola§
University of Central Florida

ABSTRACT

Firefighters rely on visual information to make tactical decisions during structure fire events. Consequently, user interface design for public safety requires a comprehensive understanding of how these visual cues impact performance in terms of expertise and domain knowledge, as well as cognitive load. However, very few studies have captured baseline measures (e.g., subjective workload) associated with the use of pre-incident planning, the phase prior to deciding how to extinguish the fire. Virtual environments and 3D modeling have been suggested as a tool that can better support the exploration of emerging technology to better support pre-incident planning practices. Sixty-four ($n = 64$) North American firefighters participated in this study. Our work focuses on providing baseline data to define effective ways to present building and hazard information to incident commanders using both current systems (e.g., 2D diagrams) and emerging technologies (e.g., 3D models, virtual reality, augmented reality).

Index Terms: Human-centered computing [Human computer interaction (HCI)]: HCI design and evaluation methods—User Studies

1 INTRODUCTION

To successfully employ fire suppression tactics, firefighters need a detailed understanding of the structure, including the location of important hazard information. However, often the information provided to firefighters, particularly incident command is lacking [1, 2]. Virtual environments and simulation have been suggested as a method to capture and measure human factors issues without interfering at the scene of an emergency. Additionally, virtyal environments can also be potentially leveraged to support firefighters training on large structures within their response areas due to the increasing availability of commercial-of-the-shelf (COTS) tools and the ability to prevent exposure to unnecessary hazardous materials [3]. This is an important concern given the health implications associated with the occupational cancer in the fire service [4]. Firefighters are more likely to be diagnosed with several different types of cancer when compared to the general population as a result of repeated exposure to carcinogens in their work environment [5]. As a result, there is an increasing interest in the use of virtual environments beyond task and procedural training, since there are broader health and safety benefits associated with the use of augmented and virtual reality for the fire service.

Virtual environments have been used to train firefighters in navigation, wayfinding, and decision-making tasks for decades [6–13]. Although these virtual environments have been explored to train these core job tasks, there is a paucity of research investigating how virtual environments could be used to support emergency preparedness efforts [14]. For example, pre-incident planning is a process in the fire service where firefighters or other relevant risk assessment personnel (e.g., inspectors, fire marshals, etc.) assess and document building information, including building construction, layout, and hazards [14]. The purpose of this pre-incident phase is to help support first responders who must make decisions within a few seconds to a few minutes by providing enough information and detail about potential hazards or building components that could impact successful response. Consequently, this phase of emergency response is a critical phase, but as previous research has demonstrated, due to sociotechnical constraints, pre-incident planning is often not prioritized enough to effectively support first responders, particularly firefighters, in the event of an emergency [14].

Sixty-four ($n = 64$) North American firefighters participated in the study. Our research is centered at the intersection of several fields, including computer science and human factors. Therefore, our work focuses on answering the following question: What is the optimal way to present this data to incident commanders using both current systems (e.g., 2D diagrams) and in future emerging technologies (e.g., 3D models, virtual reality, augmented reality)?

2 RELATED WORK

Pre-incident planning developed from a need for firefighters and first responders to understand the critical elements of a structure and to develop incident or tactical decisions based upon anticipated or forecasted fire conditions. Thus, pre-incident planning transitioned from a recommended practice to a National Fire Protection Association (NFPA) standard, requiring departments to identify elements in the environment to assist in forecasting future problems or hazards, ultimately preventing fatalities and property loss [15]. Currently, there is no recommended format or standard interface design for digital pre-incident plans; rather, each fire department is independently responsible for compliance with the NFPA 1620 standard. Because this work is multidisciplinary it involves a need to understand the problem from multiple domains and sources. Therefore, a limited discussion of the relevant constructs is posed below to emphasize the importance of this work from both a theoretical and practical perspective.

2.1 3D Modeling

Due to the increasing availability and benefits of 3D modeling, many safety-critical fields adopted 3D visualization to increase efficiency, effectiveness, and to reduce cost [16]. For example, the architecture, engineering, and construction (AEC) fields have leveraged 3D modeling for visualizing complex construction projects, like bridges or highways, leading to reductions in project time and creating fiscally responsible projects [17]. The oil and gas industries have leveraged

---

*e-mail: kate.kapalo@knights.ucf.edu
†e-mail:kevin.pfeil@unf.edu
‡e-mail:joseph.a.bonnell@phoenix.gov
§e-mail:jjl@ucf.edu

3D visualization for pipeline monitoring and for managing emergency scenarios. Additionally, in healthcare and medicine, students can leverage 3D models for more context when learning anatomy or interacting with healthcare data.

More importantly, 3D modeling provides certain advantages that 2D does not. 3D models and related simulation-based technologies can help us complete tasks that are impractical or impossible. For instance, researchers in England were able to reconstruct the Grenfell tower and demonstrate how the tragedy unfolded, which would otherwise be impossible to recreate ethically and safely [1].

2.2 The State of Pre-Incident Planning

Historically, pre-incident plans were paper-based (2D) plans contained in three-ring binders [14, 18]. As technology continues to evolve, more departments are transitioning to electronic plans, but there are currently no existing, evidence-based standards for the presentation and display of pre-incident planning information in terms of 3D modeling. Consequently, the fire service faces challenges associated with the management, retrieval, storage, and maintenance of pre-incident planning documents and information.

Due to the integration of technology and information systems, firefighters typically rely on GIS-based information to navigate to the incident scene and for information that connects different technologies like the computer-aided dispatch (CAD) system to the individual interfaces firefighters use in their vehicles [18]. Nadal-Serrano demonstrated how to leverage GIS-based technology for pre-incident planning in Madrid, Spain [19]. Additional work in this area has focused on understanding how to create templates and scalable systems for more effective digital pre-planning using additional sensors and information [2], but these efforts have largely focused on the operational aspects of the system and have not necessarily addressed the importance of collecting human performance data to demonstrate the efficacy of pre-incident plans. There have been many cases where an incorrect or incomplete PIP led to tragedy, most notably the Charleston warehouse fire in which nine firefighters were lost in the line of duty [20].

2.3 Naturalistic Decision Making

Taken together, the extant literature has important implications for this work. Namely, firefighters rely on visual information to make tactical decisions. User interface design for public safety requires a comprehensive understanding of how these visual cues impact performance in terms of expertise and domain knowledge. Therefore, studies on naturalistic decision making (NDM) lay the groundwork for research on how to develop more effective training tools and user interfaces to scaffold a firefighter's ability to identify critical cues and leverage those cues to make better decisions on the fireground, given the complexity of the operational environment [21–23]. These studies provide the background knowledge needed to support decision-making of fireground incident commanders. Since incident commanders are the ones in charge of the scene, the quality of information provided to them is critical [24]. Gasaway demonstrated through human factors methods (e.g., card sorting tasks), that situation awareness is one of the highest ranked factors in understanding successful incident management [24]. However, due to the challenges associated with information capture and dissemination during a fire incident, often times the incident commander is not able to gather enough detailed information [1, 18, 25]. Therefore, AR/VR technologies could potentially bridge this gap in information quality and availability [2, 26].

2.4 Cognitive Load and Working Memory

While there has been research focused on understanding firefighter mental workload more broadly, the majority of these studies were focused on mostly on the user interfaces and multi-modal information and did not necessarily incorporate existing building information [27]. There is hardly any research that measures subjective cognitive load associated with specific interface design to better support fireground command decision-making. To address this issue, one of the leading studies in this area measured spatial working memory [28]. By focusing on working memory, these researchers were able to demonstrate that virtual reality and 3D representations of buildings were helpful for attenuating to visual building features [29]. This work demonstrates the critical need to study how firefighters perceive spatial relationships and how AR/VR could potentially support reductions in cognitive load if firefighters are trained to understand building layouts prior to their arrival on scene.

Despite this important work, there are few studies that have captured the subjective workload from firefighter populations during a pre-incident planning task. This has resulted in a lack of existing baseline measures to determine the efficacy of different pre-incident planning formats [14]. Currently, not enough empirical research shows that pre-incident plans are effective and consequently, the use of technologies such as AR/VR is not fully supported due to a lack of existing data related to pre-incident planning use-cases [30]. This paucity of research represents a need for more empirical data to better support fireground incident commanders.

3 METHOD

The following sections outline the study participants, the study design, the specific simulated structure fire scenario, and the study deployment. We report on each of these areas to provide details regarding how the participants were able to complete the study in the height of the COVID-19 pandemic.

3.1 Participants

64 firefighters ($n = 64$) from North America (the U.S. and Canada participated in this survey). IRB approval was obtained, and data was collected from August 2020 to October 2020. Inclusion criteria required that the participant must be 18 years of age or older and all participants were required to be a current or retired firefighter. Demographic variables collected included age, years of experience, type of department, type of agency, state, and some initial questions regarding thoughts on pre-incident planning practices were captured. Data was collected across North America (including Canada) to capture the challenges and opportunities associated with pre-incident planning Participants were representative of almost all United States FEMA regions 1. Because the fire service is so diverse, one specific region's needs may not necessarily directly correlate to another. The goal of collecting this information was to better understand the current landscape of PIP practices within the fire service.

Table 1: Participant Age (years) by Study Condition

Condition	M Age	Std Dev Age	Md Age
2D	50.73	10.01	51.00
3D	49.40	10.27	48.00

3.2 Study Design

The following sections outline the study task and the experimental procedure. Our study consisted of a 2 x 1 between-subjects design with one independent variable, dimension of the PIP. Dimension – consisted of two levels, 2D and 3D. We measured cognitive load through the NASA-TLX, a validated questionnaire [31]. This test was chosen for its ease of deployment post-trial and because there have been few studies to date that have captured perceived workload during pre-incident planning tasks.

---

[1] https://forensic-architecture.org/investigation/the-grenfell-tower-fire, 2020

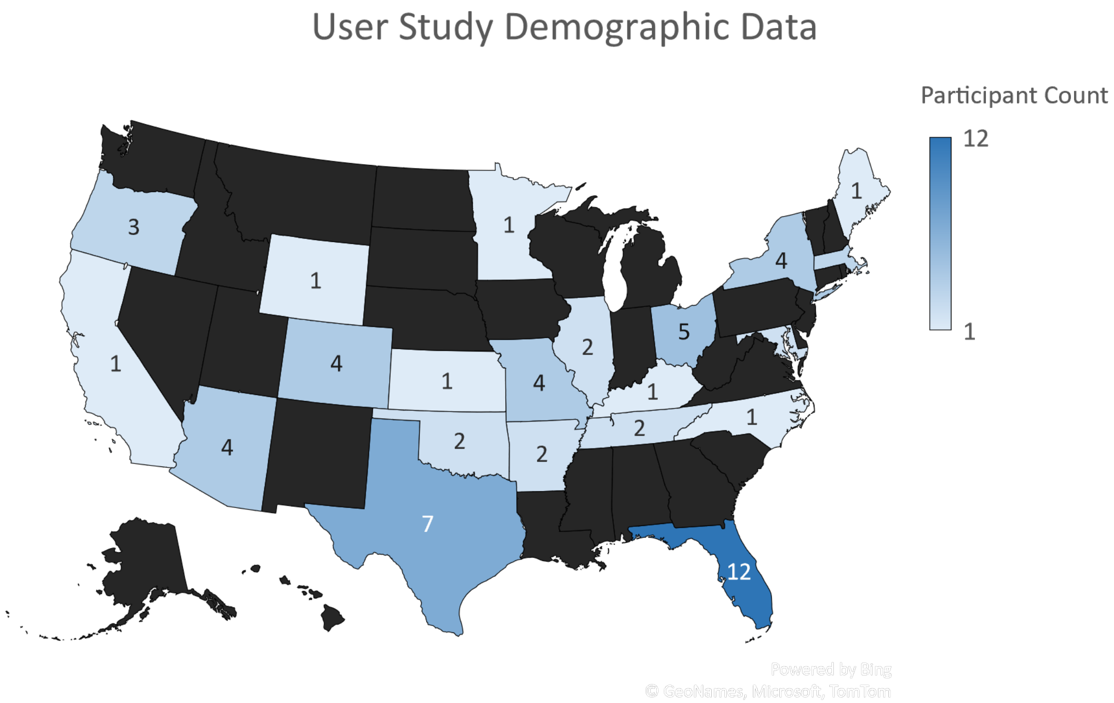


Figure 1: Participants by state (Note: Canadian participants not depicted here)

3.3 Scenario Design

The scenario for the study was developed with the support of a subject matter expert and tactical best practices. The scenario included a structure fire at a "big box," a large retail open floor plan retail store. A big box store was chosen due to some of the critical fireground factors that would influence tactical decisions. For example, the floor plan of a retail superstore such as the one in this study poses several risks to firefighters. One of the more severe challenges is air management. Because big box store floorplans are often massive, fires can be challenging to put out, and firefighters are limited in how far they can enter a structure due to the constraints of a typical self-contained breathing apparatus (SCBA). Additionally, the sprinkler system activation can cause the building to rapidly fill with cold smoke. To make the scenario realistic, we modified an existing 3D model available in the Unity Asset Store.

The scenario employed a pre-incident plan that is currently used in a large fire department located within the Southeastern region of the U.S., see Figure 2. The 3D model was customized from existing assets available in the Unity3D store [2].The goal in choosing this scenario was to compare the 2D plan (e.g., current state of PIPs) against the 3D plan. All the features present on the 2D plan were also available in 3D. Regardless of condition, participants received the same instructions and scenario details. Additionally, 3D PIP we provided in the paper can be displayed to the user akin to a world-in-miniature (WIM), which is a technique found in both AR and VR. The model also represents a scan of a building, and can be explored using egocentric AR technologies in the real word (e.g., if the incident commander was onsite) [32, 33].

In addition to the use of a current pre-incident plan, we also leveraged Blue Card [3], an internationally accredited incident command training system to ensure that the scenario met criteria that is relevant to the fire service and represents industry standards for best practices. In addition to Blue Card, approximately five company and command officers ($n = 5$) from the U.S. piloted the study and provided feedback prior to its deployment, to ensure that the design of the study was relevant to fire service members.

3.4 Survey Deployment

We leveraged Amazon Web Services to host a WebGL build of our Unity3D project, and Qualtrics to distribute our survey. We used an iframe to embed the Unity3D application into the Qualtrics survey and used client-side scripting to log which random condition (e.g., 2D or 3D) a participant received, allowing us to capture all participant clicks on the GUI for analysis. For the two conditions, the interaction method was held constant. There were several reasons for this. First, we did not intend to measure interaction methods such as gestures or touch-screen interfaces, as we specifically focused on visualization. Second, because of diversity in fire service budgets, administration, etc., each department may or may not have access to Mobile Data Terminals or Mobile Data Computers (MDTs/MDCs) that are both affordable and functional enough to retrieve electronic PIPs (e.g., tablets). Therefore, we could not hold this constant across departments, so we asked participants to complete the scenario on only a desktop or laptop computer. Finally, we could not ensure that all participants were working with the same computing equipment. Therefore, it was not possible to institute a touchscreen or a 3D user interface in this study design. We intentionally designed this GUI to be the simplest GUI that did not require extensive training to use or extra equipment to install.

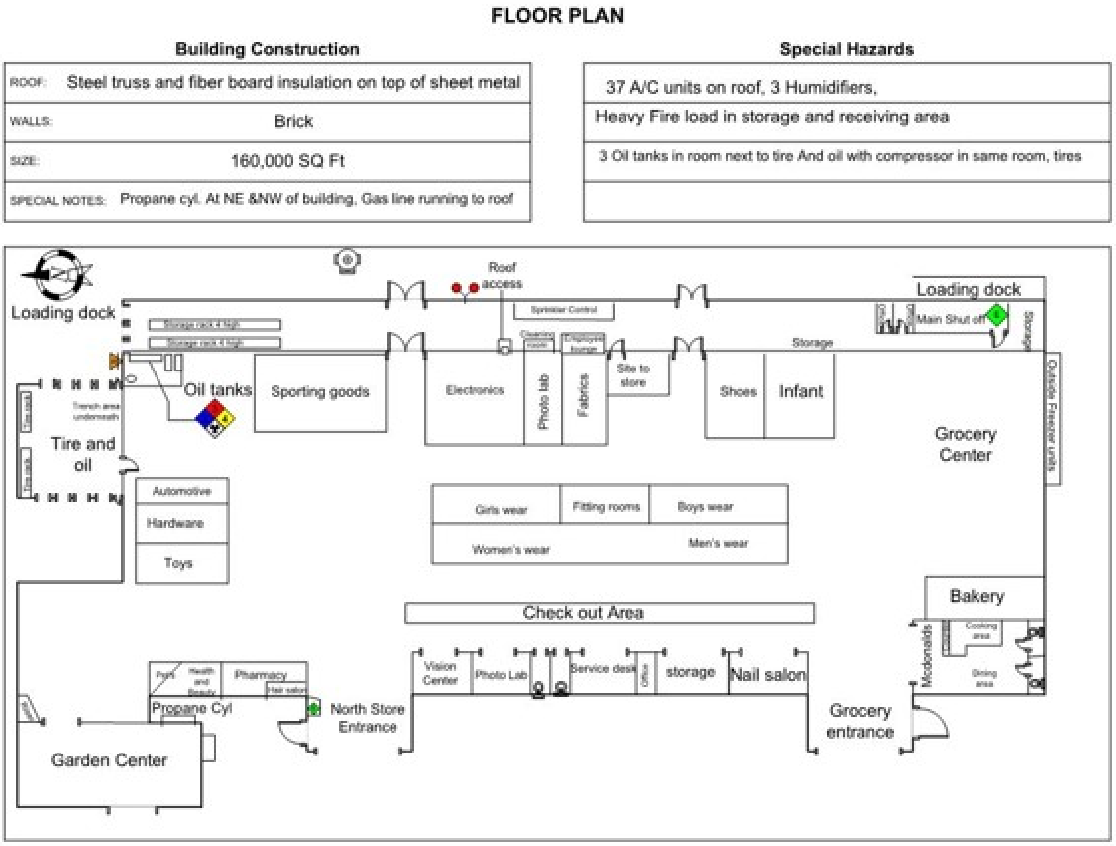
3.5 Experimental Procedure

Study participants were randomly assigned to one of the two study conditions (2D or 3D). They were given a written radio report to read and a small description of the structure. These descriptions were the same regardless of study condition. Based on the radio report and the structural information, firefighters were then asked to

---

[2] assetstore.unity.com/packages/3d/environments/small-town-america-super-store-57064

[3] https://www.bshifter.com/about$_0 1.aspx$



FLOOR PLAN

Building Construction

ROOF:	Steel truss and fiber board insulation on top of sheet metal
WALLS:	Brick
SIZE:	160,000 SQ Ft
SPECIAL NOTES:	Propane cyl. At NE &NW of building, Gas line running to roof

Special Hazards

37 A/C units on roof, 3 Humidifiers,
Heavy Fire load in storage and receiving area
3 Oil tanks in room next to tire And oil with compressor in same room, tires

Figure 2: The 2D condition (an actual pre-incident plan)

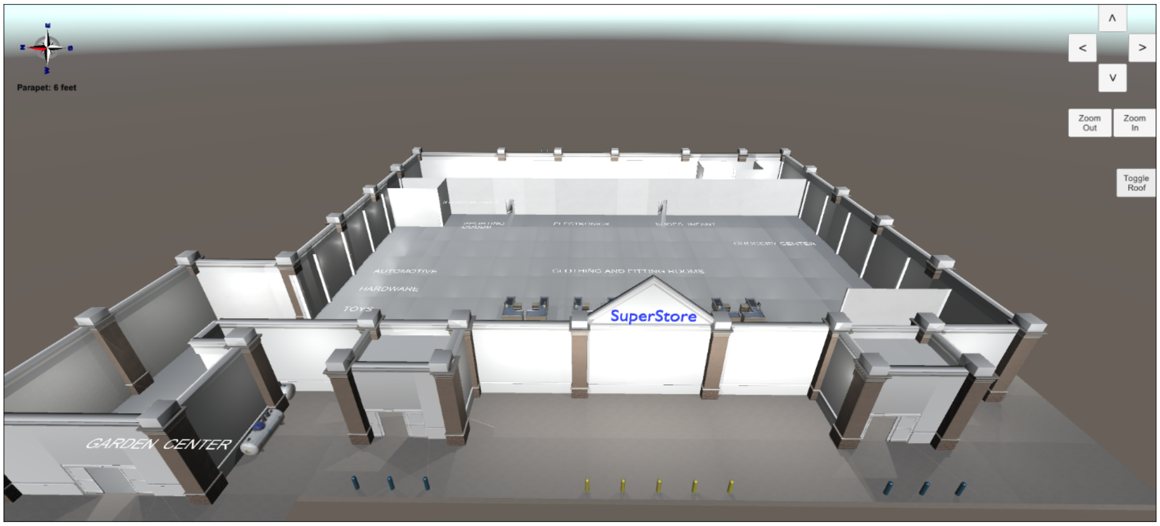


Figure 3: The 3D condition (modeled from 2D plan)

role play and complete tasks, such as correctly identifying hazards correctly in the structure, and developing an incident action plan (i.e., the tactical strategy that should be communicated out to the team). This information was collected as part of the Qualtrics survey form.

4 RESULTS

4.1 NASA-TLX Questionnaire Results

Our data violated assumptions of normality. Therefore, non-parametric tests were used to analyze the data. A Mann-Whitney U Test revealed no significant differences in NASA-TLX overall scores across the two conditions, 2D (Md = 46.67, n = 30) and 3D (Md = 45.83 , n =34), U = 9.18.50, z = -0.096, p = 0.924. Raw scores for all subscales and the overall unweighted scores are presented in the table below.

Table 2: NASA TLX Scores by Study Condition

Condition	TLX-Subscale	Raw/Unweighted
2D	Mental	55.34
	Physical	29.78
	Temporal	52.03
	Performance	62.67
	Effort	47.59
	Frustration	50.43
	Overall	25.83
3D	Mental	55.08
	Physical	26.04
	Temporal	52.86
	Performance	59.31
	Effort	49.69
	Frustration	45.19
	Overall	48.33

4.2 Qualitative Participant Feedback

In addition to the objective data, we solicited participant feedback on their overall perception of the plan and scenario. We found that most participants expressed that the 3D model helped them to better understand the fireground and the incident. For example, this is captured in the participant quote below:

> "I'm a visual learner and I've found that many folks in the fire service also seem to be. Giving firefighters "remote" eyes on the building would be a very beneficial tool in the fire service (especially for a smaller suburban area like mine where we don't often have big-box fires)." -LT, IL

In addition to comments about the 3D models, we also received some feedback on the interface design. For example, one Chief expressed frustration with the navigation tool.

> "The system provided enough information to create an IAP. The navigation buttons are a big issue."-Chief, Retired, OR

5 DISCUSSION

Overall, we found no significant differences in perceived workload between the 2D and 3D plans. Based on qualitative participant feedback, we found that firefighters overall preferred the ability to visualize the structure in 3D. This is consistent with the findings of studies in other domains such as medicine, architecture, construction, and engineering. From our data, we found that although the 3D interface did not significantly differ from 2D plans in terms of subjective workload, we were unable to draw conclusive evidence that it outperformed the standard 2D plan type that most departments currently use. It is possible that neither 2D nor 3D visualization alone can best support FGCs.

Based on the results presented in Tory et al.'s seminal work on information visualization, it may be true that 3D is best for only approximate navigation and relative positioning [34]. For measurement tasks and tasks that require the judgment of distances like apparatus placement or hose lays, 2D displays may perform better. We believe that a combination of 2D and 3D views may be more or less advantageous depending upon the occupancy type (building type) or layout of the structure [14, 30]. Future work should focus on incorporating both 2D and 3D displays to test the effectiveness of a combined approach.

5.1 Limitations

Consistent with previous research in the human factors literature, some studies suggest that raw scores are more representative of the actual workload and that the weighted subscales add confusion and frustration to the study [35, 36]. This confusion associated with the format and design of the NASA-TLX was captured in the participant quote below:

> "I was confused by your matching pairs. After reading and re reading exactly what you were asking, I still didn't have a clear idea of what I was supposed to match up or select based upon those pairs. I answered as diligently as possible but have low confidence my answers represent my true answer."-Chief, Retired, FL

We also note that there were limitations due to the ongoing COVID-19 pandemic and in-person user studies. Additionally, we had stronger responses from several states (e.g., Florida), but we did not have representation in this sample from every state. This could potentially bias the results and efforts to recruit a more representative sample are necessary. Because this study is one of the first to capture both objective data and subjective feedback specifically related to pre-incident planning, we found that cross-sectional research alone could not capture all of the nuances associated with the data.

5.2 Future Work

In future iterations of this study, we plan to extend the work and compare different types of structures and combination 2D/3D interfaces to better understand how information should be presented and displayed to FGCs. More importantly, it will be critical to add supplemental features like distances, measurements, and other relative positioning data to see if this better supports FGCs in making tactical decisions upon arrival at the incident scene. Future work should focus on streamlining the subjective measures to better support the end-users in completing the study. This study provided preliminary results that capture some of the challenges associated with measuring workload. In future work, we plan to explore more objective measures of workload, that could be triangulated to understand whether there are differences between perceived workload and actual workload. Additionally, more objective measures can help to decipher whether participants tend to conflate cognitive workload with situation awareness, a common problem in the aviation literature [35].

6 CONCLUSION

The present study provides preliminary data and measures to better address the existing research gaps related to the use 3D models for decision-making on the fireground. To date, there have been very few studies that have examined firefighters' perceived sense of cognitive load while using 3D models of a structure. Therefore, the goal of this work is to provide some baseline findings that can be leveraged to better understand how to design more effective user interfaces for firefighters operating in dangerous and unpredictable environments.

ACKNOWLEDGMENTS

The authors wish to thank the anonymous firefighters who participated in this research. This paper is dedicated to the memory of Lt. Sean Williamson and the Philadelphia Fire Department.

REFERENCES

[1] Rune Rimstad, Ove Njå, Eivind L Rake, and Geir Sverre Braut. Incident command and information flows in a large-scale emergency operation. *Journal of Contingencies and Crisis management*, 22(1):29–38, 2014.

[2] Casey C Grant, Albert Jones, Anthony Hamins, and Nelson Bryner. Realizing the vision of smart fire fighting. *IEEE Potentials*, 34(1):35–40, 2015.

[3] S Khanal, US Medasetti, M Mashal, B Savage, and R Khadka. Virtual and augmented reality in the disaster management technology: A literature review of the past 11 years. front. *Virtual Real. 3: 843195. doi: 10.3389/frvir*, 2022.

[4] Paul A Demers, David M DeMarini, Kenneth W Fent, Deborah C Glass, Johnni Hansen, Olorunfemi Adetona, Maria HG Andersen, Laura E Beane Freeman, Alberto J Caban-Martinez, Robert D Daniels, et al. Carcinogenicity of occupational exposure as a firefighter. *The Lancet Oncology*, 23(8):985–986, 2022.

[5] Hamed Jalilian, Mansour Ziaei, Elisabete Weiderpass, Corina Silvia Rueegg, Yahya Khosravi, and Kristina Kjaerheim. Cancer incidence and mortality among firefighters. *International journal of cancer*, 145(10):2639–2646, 2019.

[6] James P Bliss, Philip D Tidwell, and Michael A Guest. The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence: Teleoperators & Virtual Environments*, 6(1):73–86, 1997.

[7] Steven G Wheeler, Hendrik Engelbrecht, and Simon Hoermann. Human factors research in immersive virtual reality firefighter training: A systematic review. *Frontiers in Virtual Reality*, page 80, 2021.

[8] Jerônimo G Grandi, Zekun Cao, Mark Ogren, and Regis Kopper. Design and simulation of next-generation augmented reality user interfaces in virtual reality. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 23–29. IEEE, 2021.

[9] Randall Spain, Jason Saville, Barry Lui, Donia Slack, Edward Hill, John Holloway, Sarah Norswothy, Bradford Mott, and James Lester. Investigating a virtual reality-based emergency response scenario and intelligent user interface for first responders. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, pages 2114–2118. SAGE Publications Sage CA: Los Angeles, CA, 2020.

[10] Hendrik Engelbrecht, Robert W Lindeman, and Simon Hoermann. A swot analysis of the field of virtual reality for firefighter training. *Frontiers in Robotics and AI*, 6:101, 2019.

[11] Rory MS Clifford, Sungchul Jung, Simon Hoermann, Mark Billinghurst, and Robert W Lindeman. Creating a stressful decision making environment for aerial firefighter training in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 181–189. IEEE, 2019.

[12] Vítor Reis and Cláudia Neves. Application of virtual reality simulation in firefighter training for the development of decision-making competences. In *2019 International Symposium on Computers in Education (SIIE)*, pages 1–6. IEEE, 2019.

[13] Ruggiero Lovreglio, Xinyue Duan, Anass Rahouti, Robyn Phipps, and Daniel Nilsson. Comparing the effectiveness of fire extinguisher virtual reality and video training. *Virtual Reality*, 25(1):133–145, 2021.

[14] Katelynn A Kapalo and Joseph J LaViola Jr. Failing to plan is planning to fail: capturing the pre-incident planning needs of firefighters. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 63, pages 612–616. SAGE Publications Sage CA: Los Angeles, CA, 2019.

[15] Marty Ahrens and Ben Evarts. Fire loss in the united states during 2020. *National Fire Protection Association (NFPA)*, 2021.

[16] Steve Dübel, Martin Röhlig, Heidrun Schumann, and Matthias Trapp. 2d and 3d presentation of spatial data: A systematic review. In *2014 IEEE VIS international workshop on 3DVis (3DVis)*, pages 11–18. IEEE, 2014.

[17] Hyunjoo Kim, Kevin Orr, Zhenhua Shen, Hyounseok Moon, Kibum Ju, and Wonsik Choi. Highway alignment construction comparison using object-oriented 3d visualization modeling. *Journal of Construction Engineering and Management*, 140(10):05014008, 2014.

[18] Katelynn Amaris Kapalo, Pamela J Wisniewski, and Joseph J LaViola Jr. First in, left out: Current technological limitations from the perspective of fire engine companies. In *ISCRAM*, 2019.

[19] Jose M Nadal-Serrano. Towards very simple, yet effective on-the-go incident response preplanning: Using publicly-available gis to improve firefighters' traditional approach. In *ISCRAM*, 2010.

[20] Frank J Baker. Lessons learned a review of published investigation reports from the charleston sofa super store fire. *Professional Safety*, 54(07), 2009.

[21] Gary A Klein, Roberta Calderwood, and Anne Clinton-Cirocco. Rapid decision making on the fire ground. In *Proceedings of the human factors society annual meeting*, volume 30, pages 576–580. Sage Publications Sage CA: Los Angeles, CA, 1986.

[22] Sabrina R Cohen-Hatton and RC Honey. Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. *Journal of Experimental Psychology: Applied*, 21(4):395, 2015.

[23] Sabrina R Cohen-Hatton, Philip C Butler, and Robert C Honey. An investigation of operational decision making in situ: Incident command in the uk fire and rescue service. *Human Factors*, 57(5):793–804, 2015.

[24] Richard B Gasaway. Making intuitive decisions under stress: Understanding fireground incident command decision-making. *International fire service journal of leadership and management*, 1(1):8–18, 2007.

[25] Justin Okoli, John Watt, Gordon Weller, and William BL Wong. The role of expertise in dynamic risk assessment: A reflection of the problem-solving strategies used by experienced fireground commanders. *Risk Management*, 18(1):4–25, 2016.

[26] Christos Papakostas, Christos Troussas, Akrivi Krouska, and Cleo Sgouropoulou. Measuring user experience, usability and interactivity of a personalized mobile augmented reality training system. *Sensors*, 21(11):3888, 2021.

[27] Jan Willem Streefkerk, Wouter Vos, and Nanja Smets. Evaluating a multimodal interface for firefighting rescue tasks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 277–281. SAGE Publications Sage CA: Los Angeles, CA, 2012.

[28] Yangming Shi, John Kang, Pengxiang Xia, Oshin Tyagi, Ranjana K Mehta, and Jing Du. Spatial knowledge and firefighters' wayfinding performance: A virtual reality search and rescue experiment. *Safety science*, 139:105231, 2021.

[29] Yiqing Zhu and Nan Li. Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review. *Journal of Safety Science and Resilience*, 2(1):1–10, 2021.

[30] Katelynn A Kapalo, Patricia Bockelman, and Joseph J LaViola Jr. "sizing up" emerging technology for firefighting: Augmented reality for incident assessment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 62, pages 1464–1468. SAGE Publications Sage CA: Los Angeles, CA, 2018.

[31] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[32] Blaine Bell, Tobias Höllerer, and Steven Feiner. An annotated situation-awareness aid for augmented reality. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 213–216, 2002.

[33] Kurtis Danyluk, Barrett Ens, Bernhard Jenny, and Wesley Willett. A design space exploration of worlds in miniature. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2021.

[34] Melanie Tory, Arthur E Kirkpatrick, M Stella Atkins, and Torsten Moller. Visualization task performance with 2d, 3d, and combination displays. *IEEE transactions on visualization and computer graphics*, 12(1):2–13, 2005.

[35] Sandra G Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA: Los Angeles, CA, 2006.

[36] Ernesto A Bustamante and Randall D Spain. Measurement invariance of the nasa tlx. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 1522–1526. SAGE Publications Sage CA: Los Angeles, CA, 2008.