First Responders' Spatial Working Memory of Large-Scale Buildings: Implications of Information Format

Yangming Shi, S.M.ASCE[1]; Jing Du, Ph.D., M.ASCE[2]; Shyam Prathish Sargunam[3]; Eric D. Ragan, Ph.D.[4]; and Qi Zhu[5]

[1]Ph.D. Student, Engineering School of Sustainable Infrastructure and Environment, Univ. of Florida, 365 Weil Hall, Gainesville, FL 32611. E-mail: shiyangming@gmail.com
[2]Engineering School of Sustainable Infrastructure and Environment, Univ. of Florida, 365 Weil Hall, Gainesville, FL 32611. E-mail: du.j@ufl.edu
[3]Dept. of Visualization, Texas A&M Univ., 3137 TAMU, College Station, TX 77843. E-mail: shyam.prathish@tamu.edu
[4]Dept. of Computer and Information Science and Engineering, Univ. of Florida, Gainesville, FL 32611. E-mail: eragan@ufl.edu
[5]Ph.D. Student, Engineering School of Sustainable Infrastructure and Environment, Univ. of Florida, 365 Weil Hall, Gainesville, FL 32611. E-mail: siqiyaoyao@gmail.com

ABSTRACT

In emergency events, first responders often have to build an accurate spatial working memory of unfamiliar spaces in a short time period. This study investigates the impact of information format on first responders' short-term spatial memory of large-scale spaces via a human-subject experiment (n=63). A virtual model was created to simulate a real building on Texas A&M University campus. A total of 28 building components were modified in the virtual model. Participants were asked to review the virtual model with one of the three methods: 2D drawing, 3D model, and VR model. After the review session, the participants were sent to the real building to identify the discrepancies, and accuracy was documented as the performance measure. The results reveal that the 3D and VR groups both significantly outperformed the 2D group in spatial working memory. This study sets the foundation to further understand how instructional information affects the performance of first responders in emergency response.

INTRODUCTION

Buildings are becoming increasingly complex. By 2020, more than 20% of US buildings can be classified as smart with 50 billion connected devices and sensor (Memoori 2015). While all these technological developments will greatly improve people's indoor living experiences, a new challenge has arisen. Especially for first responders, it is becoming much easier to get disoriented or lost in a complex building. According to a report from the National Institute for Occupational Safety and Health (NIOSH), disorientation issues play a critical role in first responder fatalities in the building environment (Mora 2003). Too much environmental information burdens cognitive processing and spatial working memory. According to Vandierendonck and Szmalec (2014), spatial working memory refers to the short-term memory of identifying and storing information related to locations in the space. It is not only related to the indoor way-finding experience but is also critical to the emergency evacuation and response.

Various studies have been done to investigate factors contributing to the difference in spatial working memory in including genders (Tlauka et al. 2005), wayfinding and navigation tasks (Richardson et al. 1999), and firefighter training experience (Bliss et al. 1997). However, it remains unclear how visual information format affects people's spatial working memory in a large-scale building environment. To fill this knowledge gap, this study investigates the impact of visual information formats, including two-dimensional (2D) drawing, three-dimensional (3D) model, and virtual reality (VR), on people's short-term spatial memory of an unfamiliar building environment. To achieve this research goal, a human-subject experiment was conducted based on a real building on a university campus. Gaze and eye-tracking data was also used to test if differences in spatial working memory, if any, can be attributed to the different in attention patterns driven by various information formats.

RELATED WORK

Information Format in Architecture and Engineering: Amid the fast development of visualization techniques, different formats of visual information are used to communicate complex architecture and engineering information, such as traditional 2D drawings, building information modeling (BIM), 3D representation such as laser scanning point cloud, virtual reality and augmented reality (VR/AR). Researchers have started to investigate the influences of these visual engineering information formats in different applications. For instance, Dadi et al. (2014) studied how 2D and 3D engineering information affect workers' cognitive workload. The study did not find significant difference among three information formats (2D drawings, 3D CAD interface, and 3D printed model) in people's mental workload. The workers' mental workload is related to the background influences such as CAD experience. Sweany et al. (2016) investigated how the formats of engineering deliverables affect craft performance. They compared the impact of the 2D plan, 3D CAD model, and 3D printed model on people's craft performance. They found that the formats of engineering deliverables influenced craft performance while controlling for spatial cognition. Yu and Gero (2017) investigated how different representations (computer-generated hidden line perspective and digital photograph) of space affected the designer's cognitive load. They used eye-tracking devices to track participants' eye gaze positions and pupil size relevant to the seven designed Area of Interests (AOI) for the picture. They found that the cognitive load was related to the specific AOI, but no significant difference was found in the changes to cognitive load and different representation methods. This study adds to the previous literature by providing new evidence about how the impact of visual information formats (2D drawing, 3D model, and VR environment) on people's spatial working memory in a large-scale unfamiliar building environment.

Studies have demonstrated that virtual reality systems can benefit understanding spatial structures. The addition of visual display features such as stereoscopic visuals and head-tracked viewing have been shown to significantly improve people's ability to identify gaps or follow connection of visual geometry (Bacim et al. 2013; Ragan et al. 2013). Other research has found VR display features can improve the memory of steps in arranging 3D objects in spatial layout (Ragan et al. 2010) or remembering spatially distributed information (Ragan et al. 2012), suggesting that VR can reduce the strain on working memory to improve cognitive task performance on spatial tasks. Our study investigates if such benefits can be applied for building inspection tasks.

Eye Tracking Techniques: This study uses eye tracking to test if differences in spatial working memory can be attributed to the attention patterns driven by different information formats. The eye tracker records user's eye movement while using the system and later allows investigators to analyze the data retrospectively (Jacob and Karn 2003). In the human-computer interaction (HCI) literature, eye tracking technique is widely utilized in usability research for improving user interfaces (Goldberg and Kotval 1999), data visualization (Conati et al. 2015), and website usability reviews (Wang et al. 2014). The advantage of using eye tracking is that the integration of eye tracking technique and psychological theory can assist researchers to investigate people's perceptual and cognitive processes (Monty and Senders 2017). With these promising benefits of implementing the eye-tracking techniques, researchers have started to use eye-tracking techniques to investigate human-building interactions. For instance, Hasanzadeh et al. (2018) used a mobile eye tracker to investigate the relationship between construction workers' visual attention and their situation awareness under hazardous conditions on the construction job site. The results indicate that workers' visual attention to the hazardous areas is modulated by their level of situation awareness. Previous studies have demonstrated the effectiveness of using eye tracking in cognitive studies.

EXPERIMENT

Experiment Hardware: The Oculus Rift CV1 headset was used as the main platform for the VR condition of the experiment. Eye-tracking data was also collected for the visual context conditions. To achieve the eye tracking, a Tobii Eye Tracker 4C was mounted to the monitor for the 2D and 3D visual information conditions. The eye tracker collects users' gaze position data when they stare at the monitor. All the gaze position data and camera position data are recorded by the system with a frequency of 90 Hz. For the eye tracking in the VR environment, the system was developed based on our previous VR systems (Du et al. 2016; Du et al. 2017; Du et al. 2017; Du et al. 2018). A raycasting technique was used to record the three-axis gaze position data and camera position data 90 times per second in the virtual environment by extending an invisible ray from the center of the camera to collisions with objects in the virtual environment. This technique is widely implemented in computer graphics to study camera direction or rending pathes. To achieve these features, serval C# scripts were written based on the software developer's kit (SDK) provided by Tobii (Tobii 2018) and the application programming interface (API) provided by Unity. The participants could freely rotate their heads to view the environment with head tracking. The software used in the experiment was developed in Unity 3D (v5.6.3f1).

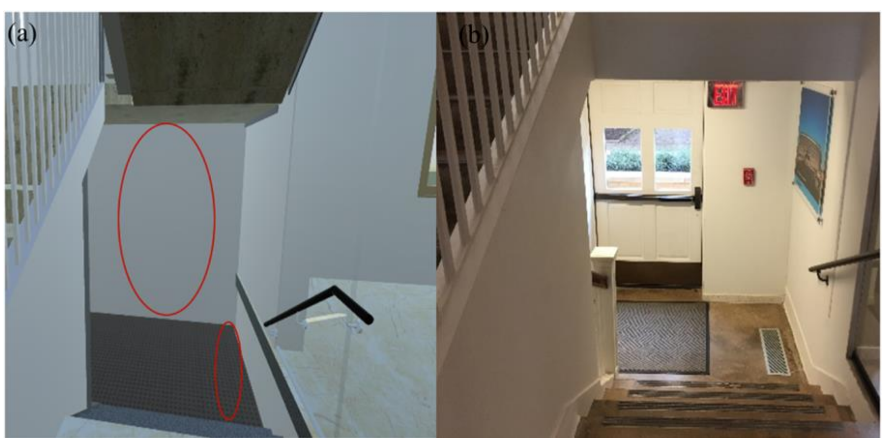

Figure 1. The discrepancies in the virtual environment and real-world building environment. (a) virtual building environment (b) real building environment

Experiment Task Design: A building inspection task was used as the main task in the experiment based on one of our previous studies (Shi et al. 2018; Shi et al. 2018). Francis Hall at Texas A&M University was selected as the testbed for the experiment. To limit the scope and time of the study to a reasonable duration for human-subjects testing, only the first floor was used. We designed 28 discrepancies that evenly distributed in the building environment which can be categorized into three types, including architectural discrepancies (e.g., windows and doors), structural discrepancies (e.g., columns), and mechanical discrepancies (e.g., air vents on the floor). For the building discrepancies, we only modified the quantity and position of the building components in the experimental building environment compared to the real-world building scenario. For instance, there were two vertical columns in one area in the real building environment, but the experimental building environment had only one column in the same area. In order to collect participants' gaze movement data in different visual information contexts, the eye-tracking system was deployed separately in different conditions. Figure 1 shows the building discrepancies in the virtual environment compared to the real-world building environment.

Experiment Procedure: One of the requirements in subject recruitment was that participants had never been to the inspection building. The experiment task required participants to find as many discrepancies as possible between the virtual models they reviewed and the real building environment. The experiment consisted of seven sessions: (1) pre-questionnaire, (2) cube test, (3) training session, (4) review session, (5) plotting session, (6) building inspection session, and (7) post-questionnaire and interview session. The pre-questionnaire session was designed to collect participants' basic demographical information such as age, gender, major, degree level, and previous game and VR experience. This session took 5 to 10 minutes. The cube test session was designed to evaluate participants' spatial cognition abilities and set their spatial cognition abilities as the baseline of their task performance. The cube comparison test developed by Educational Testing Service (ETS) was used for assessing participants' spatial cognition abilities. This session took 5 to 10 minutes.

The training session was designed for participants to get familiar with the eye tracking techniques, navigation features, and virtual environment. Participants were asked to calibrate their eye movement to the eye tracking system and the investigator ensured participants' eye movements were accurately captured by the eye tracker after serval calibration trails. The participants were also given instructions about the navigation features. For the 3D group, participants were instructed to use keyboard and mouse to navigate and change their field of view (FOV) in the virtual environment. For the VR group, participants were taught to utilize an XBox game controller to control their navigation and use their physical body rotation to control their rotation in the virtual environment. The participants were informed that there were 20 to 30 discrepancies in the building. They were also required to find as many discrepancies as they could and their performance would be compared with the other groups. The purpose of providing such information was that we wanted to motivate the participants in this experiment. This session took 5 to 10 minutes.

The review session was used for participants to review and memorize the layout and details of the building. Participants were randomly assigned to one of the three reviewing visual conditions (2D, 3D, and VR). Participants had five minutes to review and memorize the building. Five minutes was limited for the review session as some participants may feel sickness (nausea, headache, dizziness, and light headed) for 10 minutes or more in the VR environment based on our previous studies. The plotting session was designed to evaluate how much the participants remembered based on their review in the general perspective. The participants were given a paper with the outline of the building environment. Participants had to plot the discrepancies they remembered on paper. After the plotting session, participants were immediately asked to go to the real building site to identify discrepancies that were different from what they reviewed.

The starting point of the inspection was the entrance of inspection building to be consistent with the experience in the 3D and VR review conditions. The investigator of this experiment also accompanied the participants to help participants record the time point when they found each of the discrepancies. The participants were also asked to briefly describe the discovered discrepancies on a data collection form. To reduce guessing during the inspection, participants were told that there was a point penalty when they identified wrong discrepancies. A 15-minutes time limit was set to the building inspection session based on the results of preliminary studies.

At the end of the experiment, participants were asked to fill a post-questionnaire and provide comments and feedback of the experiment. The post-questionnaire was designed to collect participants' feedback of using the system including ease of control, fun, presence, sickness, and attention. The experiment procedure took approximately 60 to 90 minutes for each participant. Figure 2 shows the participants were utilizing the system to tour the building in the different visual context conditions.

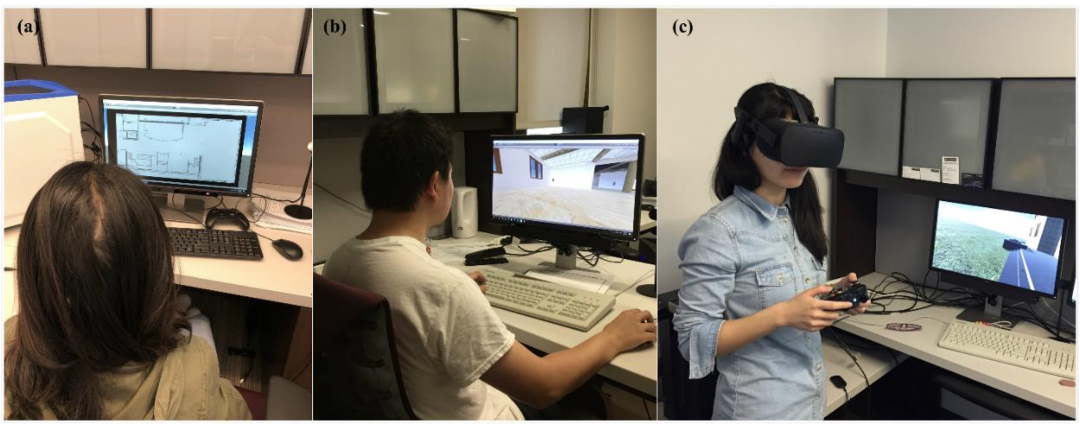

Figure 2. Participants used the developed system to memorize the building environment in different visual contexts: (a) 2D visual context (b) 3D visual context (c) VR visual context

RESULTS AND ANALYSIS

Participants: The study was approved by our institutional review board (IRB). A total of 63 participants (35 males, 28 females) took part in the experiment including 10 undergraduate students and 53 graduate students. Participants were from a variety of disciplines, and most were civil engineering students, construction management students, and architecture students. Participants' age ranged from 19 to 39, and the median age was 26. Each participant was randomly assigned to one of three groups (2D, 3D, or VR) before the experiment. Each study session took approximately 60 to 90 minutes.

Building Inspection Score: The building inspection score (BIS) is defined as the correct identifications minus the wrong identifications. The purpose of adding inspection penalty was to prevent participants from randomly guessing the discrepancies during the building inspection process. BIS was treated as a measure participants' spatial working memory performance. There were 21 participants in each of the 2D, 3D, and VR groups. Based on the Mahalanobis distance rule for outlier analysis, three outliers were removed from the results based on BIS. After removing the outliers, there were 20 participants in each of the three groups (60 total). The 2D group had M = 10.5 with SD = 4.03, the 3D group had M = 16.1 with SD = 3.28, and the VR group had M = 15.3 with SD = 4.11. Based on Shapiro-Wilk tests of normality, BIS of the three groups was normally distributed. The data meet the assumptions required for a parametric statistical test such as one-way ANOVA. The one-way ANOVA test was used to compare the BIS across three groups. We used $\alpha = 0.05$ as the threshold for significance. We found a significant difference ($p < 0.0001$) in participants' BIS across the three groups. A pairwise comparisons test for all pairs-Tukey-Kramer HSD test found differences between the 2D group and 3D group ($p < 0.0001$), between the 2D group and VR group ($p < 0.001$) and did not find a significant difference between 3D group and VR group ($p = 0.787$). Figure 3 (a) shows the results of BIS by different groups.

Visual Attention Percentage: Visual attention percentage was defined as the percentage of the aggregated attention time on all discrepancies based on participant's total review time. It represented how much time a participant paid attention to the potential building discrepancies when he/she reviewed the building environment in the general level. In order to calculate the visual attention percentage, 28 areas of interest (AOI) were defined in each group based on the building discrepancies areas. Four outliers were removed from the results based on the Mahalanobis outlier rule. The visual attention percentage in the different groups was found to be normally distributed based on the results of Shapiro-Wilk tests. Thus, a one-way ANOVA test was used to compare the visual attention percentage across different visual conditions with a significance threshold of $\alpha = 0.05$. We found a significant difference ($p < 0.001$) in the visual attention percentage of participants across visual contexts. Post hoc comparisons by Tukey-Kramer HSD testing found differences between the 2D group and 3D group ($p < 0.001$), between the 2D group and VR group ($p < 0.001$), and between the 3D group and VR group ($p = 0.006$). Figure 3 (b) shows the results of visual attention percentage. Participants in the VR group paid more attention to building discrepancies than other groups when they reviewed the building. Despite all the participants not knowing what the discrepancies were when they studied the layout of the building, the results indicated that the VR visual context helped participants to concentrate on the building environment and pay more attention to the details of building components.

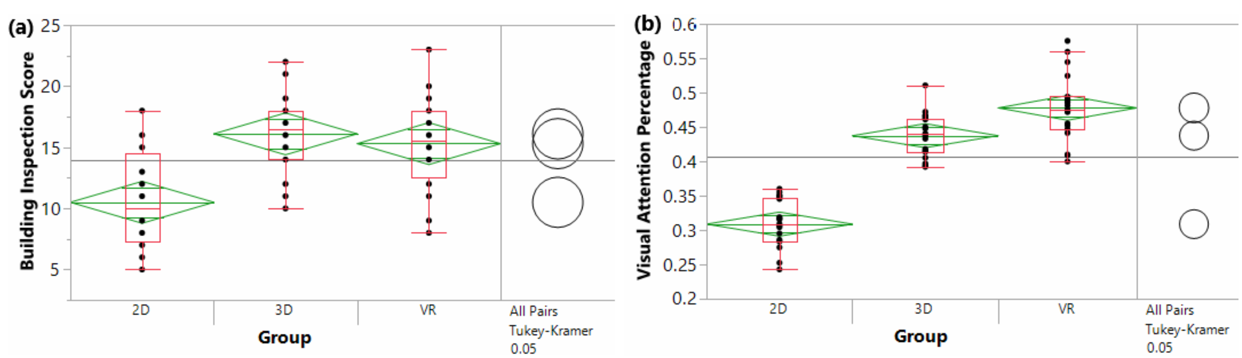

Figure 3. Statistical results of BIS and visual attention percentage across conditions.

CONCLUSION AND FUTURE STUDY

A human-subjects experiment was conducted to investigate the impact of visual information format on people's spatial working memory in the large-scale unfamiliar building environment. The results of the study reveal that the 3D and VR visual contexts both significantly outperformed the 2D format in spatial working memory, partially due to differences in attention patterns driven by different visual information formats. The 3D and VR visual contexts helped participants to concentrate on the details of the building environment compared to 2D visual context. The results demonstrate the potential value of using 3D formats to review and learn building layouts, which is essential for quality building inspection and effective emergency response in complex buildings. The results of this study will also help researchers better understand the process of forming spatial working memory in the unfamiliar building environment and help them to provide a suitable visual information context for firefighters who have to memorize the unfamiliar building environment in a short-term period. As a future research direction, more accurate eye tracker will be integrated into the system for capturing the user's gaze movement. With precise eye tracking, more complex and dynamic building scenarios could be identified and investigated. Moreover, in real-world practice, the information overload problem exists in the large and complex building environments, which should be studied in future research.

ACKNOWLEDGEMENTS

This material is based in part upon work supported by the National Science Foundation (NSF) under Grant Numbers 1761459 and 1761950, as well as National Institute of Standards and Technology (NIST) under Grant Number 60NANB18D152.

REFERENCES

Bacim, F., Ragan, E., Scerbo, S., Polys, N. F., Setareh, M., and Jones, B. D. "The effects of display fidelity, visual complexity, and task scope on spatial understanding of 3D graphs." *Proc., Proceedings of Graphics Interface 2013*, Canadian Information Processing Society, 25-32.

Bliss, J. P., Tidwell, P. D., and Guest, M. A. (1997). "The effectiveness of virtual reality for administering spatial navigation training to firefighters." *Presence: Teleoperators and Virtual Environments*, 6(1), 73-86.

Conati, C., Carenini, G., Toker, D., and Lallé, S. "Towards User-Adaptive Information Visualization." *Proc., AAAI*, 4100-4106.

Dadi, G. B., Goodrum, P. M., Taylor, T. R., and Carswell, C. M. (2014). "Cognitive workload demands using 2D and 3D spatial engineering information formats." *Journal of Construction Engineering and Management*, 140(5), 04014001.

Du, J., Shi, Y., Mei, C., Quarles, J., and Yan, W. "Communication by Interaction: A Multiplayer VR Environment for Building Walkthroughs." *Proc., Construction Research Congress 2016*, 2281-2290.

Du, J., Shi, Y., Zou, Z., and Zhao, D. (2017). "CoVR: Cloud-Based Multiuser Virtual Reality Headset System for Project Communication of Remote Users." *Journal of Construction Engineering and Management*, 144(2), 04017109.

Du, J., Zou, Z., Shi, Y., and Zhao, D. (2017). "Simultaneous Data Exchange between BIM and VR for Collaborative Decision Making." *Computing in Civil Engineering 2017*, 1-8.

Du, J., Zou, Z., Shi, Y., and Zhao, D. (2018). "Zero latency: Real-time synchronization of BIM data in virtual reality for collaborative decision-making." *Automation in Construction*, 85, 51-64.

Goldberg, J. H., and Kotval, X. P. (1999). "Computer interface evaluation using eye movements: methods and constructs." *International Journal of Industrial Ergonomics*, 24(6), 631-645.

Jacob, R. J., and Karn, K. S. (2003). "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises." *The mind's eye*, Elsevier, 573-605.

Memoori (2015). "Big Data for Smart Buildings 2015 to 2020." <https://www.memoori.com/portfolio/big-data-smart-buildings-2015-2020/>.

Monty, R. A., and Senders, J. W. (2017). *Eye movements and psychological processes*, Routledge.

Mora, W. (2003). "U.S. Firefighter Disorientation Study 1979-2001."

Ragan, E. D., Bowman, D. A., and Huber, K. J. (2012). "Supporting cognitive processing with spatial information presentations in virtual environments." *Virtual Reality*, 16(4), 301-314.

Ragan, E. D., Kopper, R., Schuchardt, P., and Bowman, D. A. (2013). "Studying the effects of stereo, head tracking, and field of regard on a small-scale spatial judgment task." *IEEE transactions on visualization and computer graphics*, 19(5), 886-896.

Ragan, E. D., Sowndararajan, A., Kopper, R., and Bowman, D. A. (2010). "The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments." *Presence: Teleoperators and Virtual Environments*, 19(6), 527-543.

Richardson, A. E., Montello, D. R., and Hegarty, M. (1999). "Spatial knowledge acquisition from maps and from navigation in real and virtual environments." *Memory & cognition*, 27(4), 741-750.

Shi, Y., Du, J., Ragan, E., Choi, K., and Ma, S. "Social Influence on Construction Safety Behaviors: A Multi-User Virtual Reality Experiment." *Proc., Construction Research Congress 2018*, 174-183.

Shi, Y., Du, J., Tang, P., and Zhao, D. "Characterizing the Role of Communications in Teams Carrying Out Building Inspection." *Proc., Construction Research Congress 2018*, 554-564.

Sweany, J., Goodrum, P., and Miller, J. (2016). "Analysis of empirical data on the effects of the format of engineering deliverables on craft performance." *Automation in Construction*, 69, 59-67.

Tlauka, M., Brolese, A., Pomeroy, D., and Hobbs, W. (2005). "Gender differences in spatial knowledge acquired through simulated exploration of a virtual shopping centre." *Journal of Environmental Psychology*, 25(1), 111-118.

Tobii (2018). "Tobii Unity SDK for Desktop." <https://developer.tobii.com/tobii-unity-sdk/>.

Vandierendonck, A., and Szmalec, A. (2014). *Spatial working memory*, Psychology Press.

Wang, Q., Yang, S., Liu, M., Cao, Z., and Ma, Q. (2014). "An eye-tracking study of website complexity from cognitive load perspective." *Decision support systems*, 62, 1-10.

Yu, R., and Gero, J. "Exploring designers' cognitive load when viewing different digital representations of spaces: A pilot study." *Proc., International Conference on Research into Design*, Springer, Singapore, 457-467.