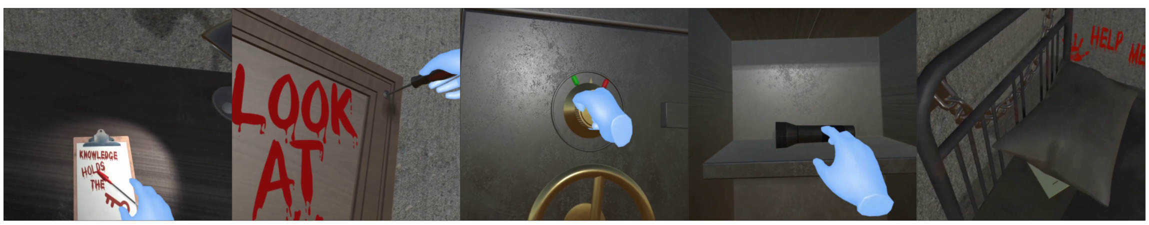
Interaction Fidelity vs User's Workload in a VR Environment: A Pilot Study

MAURIZIO MANCINI, Sapienza University of Rome, Italy
JAKE SPREADBOROUGH, School of Computer Science & IT, University College Cork, Ireland
LAURA MAYE, School of Computer Science & IT, University College Cork, Ireland
BEATRICE BIANCARDI, LTCI, Télécom Paris, Institut polytechnique de Paris, France
GIOVANNA VARNI, LTCI, Télécom Paris, Institut polytechnique de Paris, France

This paper describes a preliminary study on how Interaction Fidelity, shaped by a combination of visual, auditory and haptic modalities, impacts the user's workload. A VR escape room environment consisting of 5 puzzles to be solved in a pre-defined order is presented. Preliminary analysis shows that further investigation on the VR escape room environment could provide insights on how IF influences the user's workload depending on the type of task.

CCS Concepts: • **Software and its engineering** → Interactive games; • **Human-centered computing** → **Virtual reality**; *User studies.*

Additional Key Words and Phrases: VR, multimodality, interaction fidelity, user's workload



1 2 3 4 5

Fig. 1. A VR environment consisting of 5 puzzles: 1. opening locked drawer with the key (Padlock and Key); 2. using the screwdriver to unscrew a mirror frame from wall (Screwdriver); 3. opening the safe by entering the right combination (Safe); 4. using the flashlight to reveal hidden writings (Flashlight); 5. finding the written note hidden in the bed (Bed / Hidden Note).

1 INTRODUCTION AND STATE OF THE ART

One of the goals of virtual reality (VR) is to ensure the user's workload is not over or under loaded in problem-solving tasks. While workload has been defined in many ways, our understanding falls in line of that of Hart [7]. Hart asserts workload as the "cost of accomplishing mission requirements" [7]. Situations where the cost of maintaining performance in a task is high can result in undesirable effects with the user, such as fatigue and sickness. Thus, the goal of a designer is to ensure that workload is evenly distributed across tasks to minimise these costs.

Joint workshop on Games-Human Interaction (GHItaly21) and Multi-party Interaction in eXtended Reality (MIXR21), July 12, 2021, Bolzano, Italy

Authors' addresses: Maurizio Mancini, m.mancini@di.uniroma1.it, Sapienza University of Rome, Rome, Italy; Jake Spreadborough, School of Computer Science & IT, University College Cork, Cork, Ireland; Laura Maye, School of Computer Science & IT, University College Cork, Cork, Ireland, laura.maye@ucc.ie; Beatrice Biancardi, LTCI, Télécom Paris, Institut polytechnique de Paris, Palaiseau, France, beatrice.biancardi@telecom-paris.fr; Giovanna Varni, LTCI, Télécom Paris, Institut polytechnique de Paris, Palaiseau, France, giovanna.varni@telecom-paris.fr.

Indeed, a user's workload may be impacted by the amount of resources a person is required to complete a task, which include the number of modalities, such as visual, auditory, and haptics, implemented [12]. We are investigating workload with relation to Interaction Fidelity (IF) in VR.

According to Rogers and colleagues [11], IF can be defined as the type of modalities involved in the interaction [9], together with the "the objective degree of exactness with which a system reproduces real world interactions". In this paper, we address IF in terms of the type of the involved modalities only, keeping their degree of realism constant. So, the higher the number of types of modalities, the higher the expected IF, and vice-versa. While achieving high IF may seem the safest option, it is pertinent to investigate whether increasing the number of modalities impacts a person's performance in achieving a task, or more importantly, if employing such modalities is even needed.

To our knowledge, research on how different levels of IF and sensory feedback impact performance remains nascent. Indeed, Rogers et al. [11] and McMahan et al. [8] investigated the effect of realism degree on IF, as the visuospatial element of VR is considered the critical factor in generating immersion. Furthermore, Gall and Latoschik [6] clearly elucidate the impact that accurate multimodal feedback has on user presence. Their research suggests that the expected perceptual result of an interaction significantly correlates to a user's sense of presence. Recently, Brickler et al. [3] reveal that haptic feedback could positively support throughput in VR precision pick up and place transfer tasks. With haptics on, overall performance of the task was seen to improve. However, authors also identify that movement time increased when both haptics and audio are present. However, a decrease in movement time is seen when either only audio and haptics are present. Their study gleans the possibility of implementing audio and haptics at different stages in VR task performance.

While recent improvements of VR technology (see, e.g., exergames [1]) have enabled researchers to investigate the benefits VR may provide for various domains such as health, entertainment, training, and education [5], there is scarce work on the effect of IF in VR environments on user's workload. We deem it is important to assess this aspect of the interaction to conceive and develop novel VR interaction paradigms, enabling people to perform their everyday activities in, for example, remote settings.

In this paper, we describe a VR environment consisting of an escape room involving the solving of 5 puzzles in a pre-defined order. We aim to study how different levels of IF, shaped by different combinations of visual, auditory and haptic modalities, impact the user's workload. The exploited escape room scenario was chosen as it is a mentally and physically demanding activity consisting of everyday tasks, such as opening a locked drawer, using a screwdriver, opening a safe, using a flashlight and finding some items (see Figure 1). For example, let us consider a person leaving home to go for shopping. The person needs to find their shopping list, hidden somewhere in the kitchen; to enter their car, they need to unlock the car door; while shopping in the supermarket, they have to scan products before checking out. These activities, which are deemed fundamental for a healthy independent leaving [4], can be inhibited for some individuals in particular contexts, e.g., limited mobility due to chronic pain [1].

This study was conducted during the COVID-19 pandemic, and therefore needed to be conducted remotely. As a result, another contribution of the paper is to propose a methodology for conducting VR evaluation studies in a remote setting in which participants are self-directing themselves while carrying out the study.

2 ESCAPE ROOM

An *escape room* environment is a game in which one or more players are locked in a room setup according to a specific theme. We implemented a VR escape room environment by combining realistic visuospatial elements, with an ontologically supported narrative impetus. The opening shot of the game quickly establishes the visuospatial immersion for the player: Initially visible are the silhouette of a non-player character (NPC) watching the player from another room. The exit door and the pinpad to open the door are located to the player's left, with the first puzzle to their right. The player has five seconds to orient themselves in the scene before the initial exposition dialogue begins. Narrative impetus is continued throughout the experience as prompts and taunts are delivered by the NPC at timed intervals, while clues and flavour decals carry an aesthetic consistent with the escape room.

2.1 Puzzles

With the VR escape room environment we aim to investigate how a different number of modalities influences user's workload, based on the type of the performed task (exploration, search, manipulation, cognition). So, we designed and implemented 5 puzzles (see Figure 1) allowing the player to find the 4 digits of the pinpad and escape the room:

(1) *Padlock and Key*: the desk in the room contains three drawers, with the middle one visibly padlocked; some books on the bookshelf have to be moved out to reveal a key; once the drawer is unlocked it reveals a grabbable screwdriver and the first digit of the pinpad;
(2) *Screwdriver*: a picture frame hangs on the wall, with the words "look at me" textured on it, serving to initially draw the player's attention and notice the screws, while also serving as a callback-clue once the flashlight has been acquired (see puzzle 4); the frame hides the safe puzzle behind it;
(3) *Safe*: the safe puzzle is the focus of the pilot study presented in Section 3; the player is required to open the safe by rotating the knob through a series of actions intended to mimic a real safe; completing all the correct rotations enables rotation of the handle to open the door, which reveals the second digit of the pinpad; a blacklight flashlight is placed inside the safe;
(4) *Flashlight*: the flashlight functions as a UV-A light; the third pinpad code digit is hidden in the picture frame and can be revealed by directing the flashlight at it;
(5) *Bed / Hidden Note*: with one digit left to be found, the only option for the player is to search the remaining locations of the room; the corner of a note is visible under the pillow on the bed, giving a hint to move it out to reveal the final digit of the pinpad.

2.2 Implementation

The escape room (Figure 2 is written in C# and was designed in the Unity engine. Unity was selected as it is a powerful 3D and scene-oriented engine, with sophisticated support for mainstream VR platforms. The included package editor allowed for the OpenXR, Oculus SDK, and VRTK packages to be easily implemented in the application, with the Oculus SDK providing prefab assets that can be used to rapidly construct a player rig. Blender was used to import and adjust open-source models that were not displaying correctly in the Unity environment, or to modify certain meshes used to create the escape room geometry. Audacity was used to edit audio clips and sound effects, with these clips being acquired from open-source libraries such as freesound.org. Voicemod was used to record dialogue for the NPC, and edited via Audacity. GIMP was used to modify and create textures throughout the experience, and was particularly effective for ensuring visual consistency across the clue style. The escape room application is deployed as a reasonably small executable file (about 80 MB) that can be easily sent and played remotely with an Oculus Rift or Quest (v1 and v2) headset.

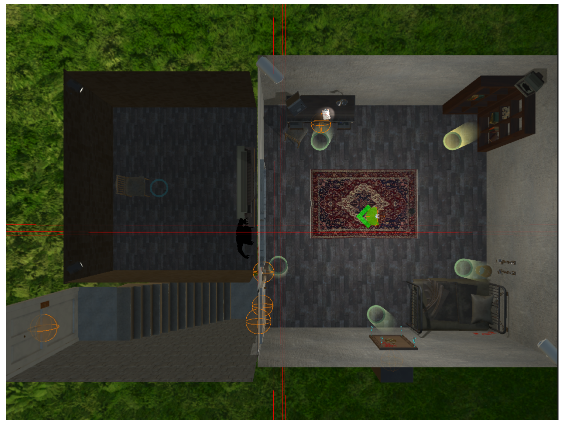


Fig. 2. Top-Down View of the Scene

3 PILOT EXPERIMENT

We designed and implemented a pilot experiment on puzzle 3 of the escape room described in the previous section: the safe opening task. So, in the pilot experiment we focus on the workload of a manipulation task (i.e., the user has to pay attention on how to use their hands in manipulating the safe knob) in relation to IF. In particular, we aim to investigate how many and which interaction modalities (independent variable) influence user's workload (dependent variable) of a user performing a manipulation task in a VR scenario (RQ1).

3.1 Conditions

To evaluate the effects of the independent variable (the modality combinations) on the dependent variable (the user's workload), four conditions using different combinations of modalities were implemented: visual only (VNN), visual-auditory (VAN), visual-haptic (VNH) and visual-auditory-haptic (VAH). According to existing works on IF (see [4]), we assume that VNN provides the lowest, VAN and VNH the intermediate and VAH the highest IF (that is, the higher the number of modalities, the higher the IF, and vice-versa).

If present, the interaction modalities look/sound/feel the exactly same throughout the 4 conditions. That is, the IF of an individual modality does not change (i.e., it is constant within that modality), while the overall IF, that takes into account all the modalities of an experiment condition as a whole, changes (i.e., it varies among the modalities).

3.2 Safe Opening Puzzle

The safe opening is the most complex puzzle developed for the escape room game (see Section 2.1). The safe has a grabbable knob and handle, and the player, to open the safe, is required to rotate the knob through a series of pre-defined positions, mimicking the opening of a real safe. Correct positions progress the puzzle, while incorrect positions reset progress. Completing all logic positions enables the rotation of the handle which, once the target value is reached, triggers the opening of the safe door.

The experiment was approved by the Ethical Committee of the University College Cork (Ireland).
An example video demonstrating the puzzle is available at: https://www.youtube.com/watch?v=uMLZEraWXis

To create a realistic safe knob, light haptic feedback is applied to the VR controller as the knob is rotated, with an array of lock-click sounds played at random in conjunction with the haptics. Additional feedback occurs when the knob enters positive and negative trigger zones. Positive zones trigger positive-click sounds and the player receives a hard vibration, while negative zones trigger negative-click sounds and an extra-hard vibration. The click-sounds were chosen to align with earcon design philosophy [2], with the negative click being a pitched-down version of the positive one.

A sophisticated logic controller script was created for the safe as it involved a large number of objects and assets to achieve the desired functionality. This script facilitates, e.g., enabling or disabling individual sensory modalities, which is effective when implementing the safe across the experiment conditions (see Section 3.1).

3.3 Participants

Participant recruitment was achieved through advertising the experiment on mailing lists and forums of researchers in VR, who could potentially have access to their personal VR equipment. In total, 8 participants (labeled P1-P8) agreed to be involved in the study.

3.4 Survey

We adopted the NASA Task Load Index (NASA-TLX) questionnaire [7] as the main part for the the experiment survey. NASA-TLX was developed to measure user's workload and it consists of six dimensions that, once combined, represent the workload experienced by people performing tasks: Mental, Physical, and Temporal Demands, Frustration, Effort, and Performance. The full version of the questionnaire is split into two parts. In the first part, each dimension is rated on a 7-point Likert scale from *1 - Low* to *7 - High* (for dimensions Mental, Physical, and Temporal Demands, Frustration, Effort) and from *1 - Good* to *7 - Poor* (for the Performance dimension only). Additionally, NASA-TLX also takes individual differences in assessing the six dimensions into account. To do that, some weights are applied in computing the overall workload score. To obtain the weights, in the second part of the questionnaire, the six dimensions are presented in pairs and the user is asked to select, for each pair, the dimension that contributes more to the overall workload of the task. The number of times each dimension is chosen determines its weight in the computation of the overall workload rating, which is equal to the sum of the weighted dimensions ratings.

The experiment survey was split into 3 main sections:

- *preliminary questions*: we asked the participant to provide informed consent to take part in the experiment and to enter anonymous demographics information (age range, gender, VR knowledge, manipulation skills);
- *NASA-TLX questions*: after playing each of the 4 conditions (VNN, VAN, VNH, VAH), the participant had to fill out the full NASA-TLX questionnaire [7]; so, this section was iterated 4 times; to avoid an influence of the order of conditions on the dependent variable, we randomised their order between participants;
- *final questions*: before the final submission, the participant was asked to fill out some qualitative questions about the user's workload experience (e.g., to specify the least/most demanding condition, by motivating their answer).

3.5 Procedure

After answering some preliminary questions about informed consent and demographics, participants were instructed to install and launch the VR application. Then, they were asked to put on the VR headset and play the first condition. Next, they filled out the NASA-TLX questions of the survey and reiterated the entire process (gameplay, followed by NASA-TLX questions) for the remaining 3 conditions. The conditions were presented in random order for each participant. Finally, they were asked to fill out the final questions that are described in Section 3.1.

3.6 Preliminary Analysis

We conducted and report here a preliminary exploration of the experimental data. The safe opening puzzle workload was computed, for each participant in each condition, as the sum of the weighted ratings provided by participants while filling out the NASA-TLX. Moreover, we computed the median of the workload and the time to open the safe within participants playing the same condition (VNN, VAN, VNH, VAH). The values are reported in Table 1.

Modalities combination	IF	$Median_w$	$Median_t$
VNN	low	0.19	29.5
VAN	mid	0.26	39.0
VNH	mid	0.29	29.0
VAH	high	0.30	45.0

Table 1. Median of users' workload score and safe opening time per condition.

Since the workload and timing scores were not normally distributed, we ran two separate Friedman tests with the modalities combination (and, consequently, the IF) as within-subjects variable and the workload score and safe opening time as dependent variables. The results did not indicate significant differences between the conditions: $\chi^2(3) = 1.412, , W = 0.06, p = 0.7$ (workload); $\chi^2(3) = 1.95, W = 0.08, p = 0.583$ (safe opening time). These results were not surprising, due to the very low sample size (8 participants). However, as we said above, we are still collecting participants, and we expect significant differences between conditions to emerge as the number of participants will increase.

In the third part of the survey we asked participants to answer a set of qualitative questions about the task workload. Below, we reveal how the participants responded to the scenarios across two questions:

- *What elements were the most distracting?* Ranked in order of frequency chosen, the most distracting conditions were VAH (4), VNH (2), VAN, (1), VNN (1). Visual-only ranked as one of the least distracting. When prompted to justify their choice, participants who chose VAH, VNH and VAN were distracted by the diversity of feedback. For P6, the simultaneous presence of audio and haptics resulted in them making *"some mistakes in rotating the knob"*. For P3, P4, P5 and P6, the ticking sound of the timer heightened anticipation, which impacted their focus on the task. As P3 highlighted: "the ticking of the timer made me nervous for like 3-5 seconds until I realized there is enough time to complete the task". P3 also commented on the realism of the haptics as also being a potential distraction: "it just doesn't feel like gripping in the real world" (P3).
- *Which condition required the most/least effort?* For 5 participants, the last condition they chose was ranked as the least effort. As P1 emphasised, they were *"already familiar with the safe logic"*. However, for P4, who chose VNH, the reason was because *"there was the haptic feedback only"*. P4 indicated that the sound of the timer was distracting, which might appear that removing this stimulus helped the user to focus on the task at hand.

4 CONCLUSION

This paper presents a VR escape room environment consisting of 5 puzzles that must be completed in a given order. We aim to study if and how different levels of IF, shaped as modalities combinations, influence the user's workload depending on the type of task. In the pilot experiment described in the paper, we focus on one of these puzzles, the safe opening one.

Despite their number being relatively low, we collected some qualitative impressions from participants, carrying out a preliminary analysis of the workload and time to open the safe.

The main contributions of the paper are: (1) to conduct a first investigation on the workload of performing different types of task in a VR environment; (2) to define a methodology to carry out VR studies in a remote setting.

Further research on the combination of modalities, e.g., sequential vs parallel [10] could be completed to identify how this impacts a person's workload.

REFERENCES

[1] M. S. H. Aung, A. Singh, S. L. Lim, A. C de C Williams, P. Watson, and N. Bianchi-Berthouze. 2013. Automatic recognition of protective behaviour in chronic pain rehabilitation. ACM.

[2] S. A. Brewster, P. C. Wright, and A. D. N. Edwards. 1995. Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of HCI*, Vol. 95. 155–159.

[3] D. Brickler, R. J. Teather, A. T. Duchowski, and S. V. Babu. 2020. A Fitts' Law Evaluation of Visuo-Haptic Fidelity and Sensory Mismatch on User Performance in a Near-Field Disc Transfer Task in Virtual Reality. *ACM Trans. Appl. Percept.* 17, 4, Article 15 (Dec. 2020), 20 pages. https://doi.org/10.1145/3419986

[4] M. Cassarino and A. Setti. 2015. Environment as 'Brain Training': A review of geographical and physical environmental influences on cognitive ageing. *Ageing research reviews* 23 (2015), 167–182.

[5] J. Collins, H. Regenbrecht, T. Langlotz, Y. S. Can, C. Ersoy, and R. Butson. 2019. Measuring cognitive load and insight: A methodology exemplified in a virtual reality learning context. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 351–362.

[6] D. Gall and M. E. Latoschik. 2018. The Effect of Haptic Prediction Accuracy on Presence. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 73–80. https://doi.org/10.1109/VR.2018.8446153

[7] S. G. Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[8] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady. 2012. Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game. *IEEE Transactions on Visualization and Computer Graphics* 18, 4 (2012), 626–633. https://doi.org/10.1109/TVCG.2012.43

[9] M. Nabiyouni, A. Saktheeswaran, D. A. Bowman, and A. Karanth. 2015. Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 3–10.

[10] L. Nigay and J. Coutaz. 1993. A design space for multimodal systems: concurrent processing and data fusion. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. 172–178.

[11] K. Rogers, J. Funke, J. Frommel, S. Stamm, and M. Weber. 2019. Exploring interaction fidelity in virtual reality: Object manipulation and whole-body movements. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14.

[12] Christopher D. Wickens. 2002. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science* 3, 2 (2002), 159–177. https://doi.org/10.1080/14639220210123806 arXiv:https://doi.org/10.1080/14639220210123806