

Diegetic Tool Management in a Virtual Reality Training Simulation

Patrick Dickinson*
University of Lincoln

Andrew Cardwell†
University of Lincoln

Adrian Parke‡
University of West Scotland

Kathrin Gerling§
KU Leuven

John Murray¶
University of Hull

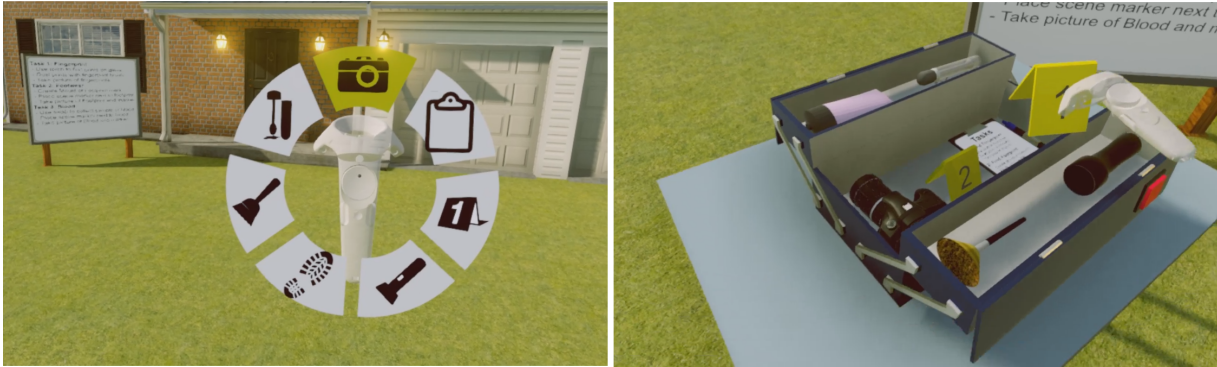


Figure 1: Spatial menu (left) and diegetic toolbox interface (right)

ABSTRACT

Researchers and developers have suggested that the use of diegetic interfaces can enhance users' sense of presence and immersion in virtual reality (VR) applications. While concepts of diegetic interfaces in VR are analogous to those seen on 2D displays, little work has considered how they might integrate with the movement-based controllers commonly used in consumer VR systems, to create higher fidelity interactions. In this paper we present a study ($N = 58$) in which we compare participants' experiences of diegetic and non-diegetic tool management interfaces, in a prototype VR crime scene investigation (CSI) training application. Contrary to expectations, we do not find evidence that participants' sense of presence is elevated when using the diegetic interface; however, we suggest that this may be due to reported higher levels of perceived workload, which can act to degrade user experience and engagement. We conclude by discussing the relationship between diegetic interface design and interaction fidelity, and highlighting trade-offs between fidelity, engagement, and learning outcomes in VR training applications.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Interaction design—Interaction design process and methods—User interface design; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

1 INTRODUCTION

Virtual reality (VR) is an engaging platform which is receiving increasing attention due to the recent prevalence of low-cost, high-

fidelity consumer systems. The experience of users is characterised by a strong sense of *presence* [38], which has been leveraged by developers and researchers working in a number of contexts, including applications to training: users can be placed into simulated scenarios which would otherwise be difficult to replicate in real life, such as aircraft evacuations [4] or natural disasters [15]. The importance of presence in learning processes is still an active area of research, and there is some evidence that perceived presence may correlate positively with enhanced training outcomes, and overall engagement (e.g. [6, 20, 39]), and is therefore a desirable characteristic.

The term *diegetic interface* describes a user interface for a simulated world (either in VR or on a 2D screen), where controls or information appear as part of the simulated world itself, rather than on a separate menu system or screen overlay. Examples from the literature include Salomoni et al. [30], who created a diegetic menu interface for a VR game by placing it on a tablet device that appeared as an object within the game world. Similarly, Peacocke et al. [23] created a diegetic ammunition counter for a shooter game running on a 2D display by placing the counter directly on the player's gun. Previous research suggests that diegetic user interfaces can contribute positively to user experience, engagement, and sense of immersion (e.g. [5, 12, 30]). However, such work has mainly focussed on users of traditional 2D displays. Few existing studies have investigated diegetic interfaces in VR applications, and none have done so in the context of training simulations.

In this paper we present a study ($N = 58$) in which we investigate the affects of diegetic interface elements on participants' sense of presence and workload, in a prototype VR training simulation which has been designed to help teach the procedural aspects of crime scene investigation (CSI). Real-world CSI procedures involve the transportation, selection and manipulation of a variety of tools. In our study, participants use two version of the simulation: one in which virtual tools are selected using a (non-diegetic) spatial menu, typical of those commonly seen in VR applications, and one in which the user carries them in a (diegetic) toolbox.

Our diegetic interface requires users to replicate bending and reaching movements which more closely replicates real-world manipulation of tools, creating higher levels of interaction fidelity than the non-diegetic menu. Higher fidelity systems are generally con-

*e-mail: pdickinson@lincoln.ac.uk

†e-mail: acardwell@lincoln.ac.uk

‡e-mail: adrian.parke@uws.ac.uk

§e-mail: kathrin.gerling@kuleuven.be

¶e-mail: john.murray@hull.ac.uk

sidered desirable, as they may carry additional learning benefits [1]; however, our interest is focussed on user experience, and previous work has reported both positive and negative aspects in this respect. For example, McMahan et al. [18] reported higher levels of presence when playing a game with a higher fidelity interface, whilst Rogers et al. [28] reported that some users found that a high fidelity interface which utilises extensive body movement to be uncomfortable and challenging. Thus, our work contributes to an understanding of the trade-offs involved in designing user interfaces for training systems, and other comparable VR experiences, from a user experience perspective.

2 RELATED WORK

In this section we review previous related work. We begin by considering users' sense of presence, and its relevance to training applications. We then proceed to consider work in the connected areas of interaction fidelity and diegetic user interfaces (both in VR and in applications for traditional 2D displays), which we use to motivate our approach in a concluding summary.

2.1 Presence and Training Applications in VR

Presence refers to the mediated sense of being present in another virtual (or real) place [38], and is a strong characteristic of immersive VR simulations. This phenomenon has been extensively studied. For example, Slater and Wilbur [36] describe it as “*a state of consciousness, the (psychological) sense of being in the virtual environment.*”. The experience of mediated presence has been deconstructed: Slater [35] defined the terms *place illusion*, which refers to the said sense of “being there”, and *plausibility illusion*, referring to the sense that what is happening in a simulation is real. A number of factors are known to affect users' sense of presence in VR, including visual realism [35]. Regenbrecht et al. [26] showed that users' sense of presence and “realness” were also enhanced by the possibility of interactions within a virtual environment.

VR has been used and studied extensively for applications in training and education (a recent review is presented by Jensen and Konradsen [13]). The relationship between system immersion, users' sense of presence, and training effectiveness have been investigated; for example, Chowdhury et al. [6] showed that, for a VR wheelchair simulator, presence correlates positively with learning recall. Stevens et al. [39] showed a correlation between presence and performance, for a military gunnery task. Similarly, Ragan et al. [25] showed that higher system immersion can positively affect recall in spatial memorisation tasks. However, Buttussi and Chittaro [4] found that although higher display fidelity produced higher levels of presence in a VR airline safety training simulator, it did not improve learning outcomes. Thus, whilst a direct relationship between presence and learning outcomes is not fully established, there is evidence that higher levels of presence can have a positive effect (as well as facilitating generally more engaging experiences) and is thus a desirable characteristic of VR learning experiences.

2.2 Fidelity and Movement-based Interfaces

Fidelity describes the degree to which a simulation replicates the real-world, and in VR is considered to be orthogonal to immersion [34]. It is also considered an important attribute of training simulations: Alexander et al. [1], distinguished *physical*, *functional* and *psychological fidelity* as distinct components, and concluded that whilst high fidelity may support better learning outcomes, and is therefore desirable, “*the appropriate level of fidelity for a training system is dependent on the skills or behaviours that are to be trained*”. There is some evidence that higher levels of fidelity can enhance users' sense of presence, in particular contexts (e.g. [19], [40]), although relatively little work has examined this relationship.

The concept of *interaction fidelity* (e.g. [19]) relates to the naturalness or realism of a user interface while engaged in a simulation.

This could include, for example, the control of tools in medical simulations (where high fidelity is regarded as important) [27]. Relevant recent work in VR includes Rogers et al. [28], who investigated interaction fidelity in object manipulation and full-body movement tasks, for VR games. They found that, for object manipulation, higher interaction fidelity increased players' sense of presence and game engagement. However, more complex full body movements were perceived negatively by some users, who found them challenging and awkward, suggesting some level of trade off between engagement and usability. Speicher et al. [37] compared isomorphic (naturalistic) and non-isomorphic methods for manipulating objects in a VR shopping application. The non-isomorphic methods offered overall better user experience and lower perceived workload, despite being less naturalistic, suggesting a similar trade off.

While movement-based controls are now standard in commercial VR systems, researchers studying games on traditional 2D displays have also explored natural interfaces using similar movement controls. For example, McGloin et al. [16] studied controller naturalness for a tennis game. Players using a movement-based interface to hit the ball, replicating the physicality of real tennis, experienced an enhanced sense of presence and enjoyment compared with using a traditional controller. Similarly, Shafer et al. [33] also found elevated levels of presence and enjoyment for a range of games when using natural movement interfaces. Birk and Mandryk [3] found that movement-based controls for a flying game elevated player experience on a number of measures including affective response, immersion, and intrinsic motivation.

Thus, there is evidence that higher-fidelity, more natural, movement-based interfaces can increase users' sense of immersion, engagement, and presence (particularly in games) both in VR and with traditional 2D displays. However, interfaces requiring more complex or demanding physical movements may be perceived as awkward, degraded users' experiences in some instances.

2.3 Diegetic Interfaces and User Experience

The term *diegesis* originates in literary and film theory, and was adapted to describe video game components by Galloway [9]. In this context, Galloway describes it as “*the game's total world of narrative actions*”, as opposed to non-diegetic elements which are “*outside the portion of the apparatus that constitutes a pretend world of character and story*” (for example, a head-up display or screen overlay). The term *quasi-diegetic* is also sometimes used (e.g. [30]) to describe game elements which occupy the same 3D space as the game world, but are not part of the game world narrative. Such structures are common in VR simulations; for example, spatial menu systems are often used ([7,21,31]).

Thus far, the study of diegetic interface elements has mainly focused on games, using traditional 2D displays. Iacovides et al. [11] investigated the effects of removing non-diegetic interface elements on immersion, for players of a first-person shooter (FPS) game. Experienced players reported an increased sense of cognitive involvement and control. Similarly, Peacocke et al. [22] found that players of a shooter game experienced increased sense of immersion, and better performance, when using a diegetic ammunition counter, rather than a head-up display. Extending this to other types of in-game information [23], they concluded that player performance can be (but is not always) improved by diegetic information displays. In contrast, Pfister and Ghellal [24] found higher levels of immersion for a platform game when using non-diegetic interface elements. These contradictory results suggest that although diegetic interfaces have the capacity to engage players, other design considerations are also pertinent and sometimes conflicting.

Very little work has thus far studied diegetic interfaces in VR; however, Salomoni et al. [29,30] conducted a study which compared diegetic and non-diegetic interfaces in a VR shooter game. Whilst they used a relatively small number of participants (ten), and con-

textualised their work to a game experience, their results suggest that players experienced a heightened sense of presence while using the diegetic version. This supports the notion that diegetic representation of interfaces might enhance user experience in VR games, analogous to results for games played on 2D displays ([11, 22]).

2.4 Summary

Modern VR applications typically employ interfaces which combine movement based actions (e.g. locomotion [40]), with non-diegetic spatial menus for functions such as item selection [14, 21, 31]. However, previous research has shown that use of diegetic user interfaces can enhance users' sense of absorption and engagement with systems using 2D displays (e.g. [11, 23]), and similarly that high fidelity movement-based controls can enhance users' sense of presence in games [16, 33].

In VR systems, with movement-based controls, concepts of diegetic interfaces can overlap with those of interaction fidelity (e.g. the object manipulation task described by Rogers et al. [28]). However, this is not necessarily the case. For example, the fantasy nature of the game described by Salomoni et al. [30] meant that they did not consider questions of interaction fidelity: moving game menus onto an in-game tablet device creates a diegetic interface, but does not replicate realistic real-world operations.

Given that high fidelity is considered an important feature of training applications [1], and that very little work has so far explored diegetic interfaces in VR, we are motivated to investigate how diegetic interfaces which emphasise interaction fidelity might affect user experiences in VR training applications. In particular, we are interested in whether such interfaces positively affect users' sense of presence, as compared to spatial menu systems, and could therefore be used to create more engaging experiences.

3 STUDY OVERVIEW

We designed a two-condition within-subjects study to compare participants' experiences of using diegetic and non-diegetic interfaces in a CSI training simulator. The simulator ran on a consumer VR system, and participants performed a series of pre-defined tasks in both conditions. Tool management and selection is a central feature of the simulation, and users need to switch tools frequently while performing tasks. We therefore chose this interface component as the independent variable for our study. The non-diegetic condition provides a spatial menu system operated using a track pad on the controller (described in Section 3.1.1). In the diegetic version, tools are kept in a toolbox which the user can carry, open, and close (described in Section 3.1.2). This mirrors real-life, where investigators typically use a box to organise and transport their tools in the field. The toolbox interface replaced all menu functionality in the simulation. We compared participants' experiences of these two conditions, and address the following research question:

- **RQ1: Does the diegetic tool management method affect users' sense of presence?** Whilst previous literature suggests that diegetic interfaces may increase users' sense of immersion and presence, no existing work has empirically assessed the affects on users' sense of presence in VR, in the context of training.
- **RQ2: What is the affect of the tool management interface on perceived workload and performance?** Diegetic and higher-fidelity interface elements may have negative affects on usability or performance [28, 37], so we also wish to explore this in our study conditions.

3.1 The Crime Scene Training Simulator

We adapted a pre-existing prototype VR CSI training simulator to conduct our study. The simulator was developed as a prototype



Figure 2: Inside the simulated house

training tool, in conjunction with staff from the School of Chemistry at the University of Lincoln, and was designed to enable students to practice aspects of CSI fieldwork, which would normally be conducted in a real-world reconstructed crime scene house.

The simulator runs on the HTC Vive virtual reality platform, and the environment comprises a fully modelled house, and front garden where the user starts their experience. The user does not have an avatar, but can see and use the tracked controllers to manipulate objects within the simulation. The floor area of the crime scene is large, and so locomotion is achieved using a combination of teleporting (using one of the controllers to select a new location), and walking (to achieve natural and fine-grained control of position and orientation). The teleporting mechanic is typical of that seen in commercial VR applications. Objects in the house are interactive, and can be moved and picked up using the controllers. No other human characters appear in the simulation. Figure 2 shows screen shots of rooms inside the house.

The simulation provides the user with a variety of typical CSI tools, which can be held and manipulated using the controllers. For our study we made use of a reduced number of commonly used tools (black-light torch, fingerprint brush, footprint gel mould, scene markers, blood swab, and camera), to make the simulator more accessible to users with little or no experience of CSI techniques, and to help us design a clear layout for the toolbox. We designed a fictitious scenario for our study which was representative of a typical incident attended by investigators, and would require the use of the available tools. The house was set up to represent a scene following a minor break-in: the rear glass patio door of the house had been forced open and broken by a hypothetical intruder (Figure 3), and there was a small amount of blood on the kitchen work surface where the intruder had accidentally injured themselves during the break-in.

For our study we compared two versions of the CSI simulation. The first had a radial tool selection menu, which is non-diegetic in the sense that it does not represent a part of the simulation narrative, nor does it exist as an object within the physically simulated world (e.g. it cannot interact physically with other objects). This menu is the user's entry point for tool selection, and so the central part of the



Figure 3: Broken Rear Door

user interface (aside from locomotion). In the second version, the radial menu was replaced with a diegetic toolbox interface, which is used to select and manage tools. The toolbox exists as an object in the simulated world, and interacts with the environment, and other objects. Both interfaces are described in more detail below.

3.1.1 Radial Menu Interface (Non-diegetic)

We decided to use a radial menu design, centred on the controller. This places the menu next to the current/selected tool, and the radial design naturally aligns with the circular design of the controller thumb-pad. We designed the menu operation through iterative prototyping, and based on our experience of using other VR applications. The radial selector is opened by lightly placing the thumb onto the pad. The menu, shown in Figure 1, appears, and allows the user to then highlight a tool by moving their thumb into the corresponding segment of the pad circle: haptic feedback in the form of short vibration is given when a new tool is highlighted. Depressing (clicking) the pad selects the highlighted tool, which then appears, attached to the controller. When the user lifts their thumb from the pad, the menu closes. The user may deselect a tool by using the menu to select the currently selected tool again, which causes it to disappear, or simply select a new tool. Note that tools appear and disappear on demand, and so are distinct from other objects in the simulated world.

3.1.2 Toolbox Interface (Diegetic)

The toolbox object exists within the simulated world as a diegetic object. The user can pick the toolbox up, carry it around, and place it anywhere in the world (e.g. on the floor). The box opens as animated expandable layers, reminiscent of a real toolbox mechanism, also shown in Figure 1. The expanding layers provided more surface area for laying out the tools, and is a familiar design for novice users. The user opens and closes the toolbox by touching a button on the top panel. Although this is less realistic, we felt that a manual opening mechanism would be cumbersome, due to the articulated nature of the box.

Once open, the user can take tools out of the box and manipulate them using the controllers. The user places the controller so that it is touching the object, and then presses the grip buttons (replicating a real grasping action): the tool then becomes attached to the controller. The user manipulates tools in the same way as other world objects, and can put them down anywhere in the environment (which is not possible when using the radial menu), reinforcing their diegetic nature. The user transports tools by placing them back into the box, closing it, and then carrying it to a new location. Tools have fixed positions within the box, and snap into those positions when placed back inside. This enables the user to easily pick tools from the box with the controllers. Aside from this, the tools themselves are operated in the same way as in the radial menu condition. Referring to McMahan's FIFA framework for analysing interaction fidelity

[17], elements of the interface exhibit moderate Biomechanical (kinematic) Symmetry with real world tool manipulation, such as bending to pick up the box, and reaching to collect tools, so we consider that the diegetic version is characterised by higher levels of interaction fidelity.

3.2 Hypotheses

With reference to our research questions RQ1 and RQ2, and the CSI simulator application described in Section 3.1, we make the following specific hypotheses:

- **H1:** Based on previous work ([16,28,30,33]), we hypothesise that participants will report a higher sense of presence while using the toolbox interface.
- **H2:** Participants will experience a higher sense of workload when using the toolbox, and also record longer task completion times, due to higher levels of physical movement and associated sensory-motor complexity.

4 METHOD

We recruited 58 participants (M = 28, F = 30), mainly from the campus of the University of Lincoln. All participants were over the age of 18 (Mean = 23.79, SD = 7.87), and 29 had previously used VR equipment. Participants were not pre-selected for subject area knowledge: this enabled us to access a larger participant pool, which we consider representative of novice users. We consider this an appropriate sample for our study, as our research questions are concerned with participants' experience of the user interface, rather than learning gains. However, four participants reported that they were either studying, or had studied, forensic science. Our study procedure was approved by the ethics committee of the College of Science, at the University of Lincoln.

4.1 Measures

We used the following self-report measures to address our hypotheses. Participants' sense of presence (H1) during the conditions was determined using the IGroup Presence Questionnaire (IPQ) [32]. We chose this measure over other self-report measures of presence as it is commonly used, and comprises relevant sub-scales, including *realism*, which we considered might differentiate between the diegetic and non-diegetic experiences. To determine participants' sense of workload (H2), we used the NASA-TLX [10] self-report questionnaire, along with performance metrics which recorded the time taken to complete the tasks in each condition. Finally, we conducted a brief semi-structured interview with each participant on conclusion of both conditions, during which we asked how they felt during the experiences, how they felt using each interface, and for suggested improvements to each interface.

4.2 Procedure

Participants were first screened for uncorrected vision or other conditions which might adversely affect their ability to safely use the VR equipment. They then read a participant information sheet which described the study and procedure, and were given the opportunity to ask questions. They then provided written consent, completed a brief demographics questionnaire, and proceeded to training.

4.2.1 Participant Training

Before undertaking the conditions, participants were shown how to safely use the headset and controllers. They then undertook a guided tutorial. This was conducted using a modified version of the simulation, where they were located outside of the crime scene house (the study conditions took place inside the house).

In the tutorial, participants were shown how to move (teleport), and how to use and manipulate each of the tools used in the conditions. They were shown how to open and close the toolbox, how to

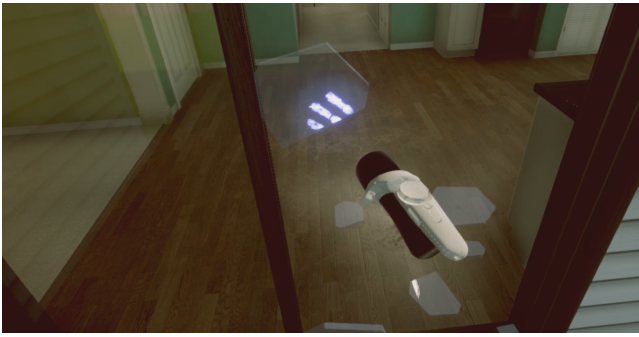


Figure 4: Using the black-light torch to identify the fingerprint (radial menu).

put tools into the box, take them out again, and how to carry the box. Participants were then shown how to use the radial menu interface to select each tool from the menu.

Finally, after completing the training, participants were shown a video describing how to complete the sequences of tasks required during the study conditions. The video was created from a recording of one of researchers performing the tasks, but was edited so that interface elements were not shown, in order to avoid biasing participants. The video showed the location of evidence in the house, which tools to use, and the order that tasks should be completed in. When participants indicated that they understood the tasks and how to operate the simulation, they proceeded to the study conditions.

4.2.2 Tasks and Conditions

In both conditions, participants completed the same predefined sequence of tasks:

- Entering through the front door, they approached the rear (open) patio door of the house, which was broken. Facing the patio door, the torch was used to locate a fingerprint on the glass, which was then dusted, and photographed using the camera.
- Moulding gel was then placed on a footprint by the door, a scene marker placed next to it, and a photograph taken of the footprint.
- Finally, participants moved to the kitchen, where a blood splash was located on the work surface. A swab was used to sample the blood splash, then a second scene marker placed, before another photograph was taken.

The sequence takes approximately three minutes to complete, and comprises a minimum of nine tool selection operations. It was designed to be representative of an actual CSI learning example, while remaining easy to remember, and simple enough for inexperienced users to perform using the reduced tool set. Participants were also provided with a description of the tasks on a clipboard tool, should they need to refer to it while performing the sequence. Figures 4 and 5 show screen captures from the tasks.

To account for learning effects, half the participants used the radial menu interface first, and the other half used the toolbox first. After each condition, participants completed the IPQ and NASA-TLX. A log of participant operations, including timing data, was automatically recorded by the simulation and stored in a text file. This was used to calculate task completion times. After completing both conditions, participants undertook the semi-structured interview, which was audio recorded and then later transcribed. Finally, participants were provided with a study debrief sheet to take away.



Figure 5: Using the camera to photograph the footprint (toolbox).

Table 1: IPQ Summary Statistics for Toolbox and Radial Menu conditions

Subscale	Toolbox		Radial		p-value
	Mean	SD	Mean	SD	
Spatial Presence	22.81	3.50	22.33	3.73	0.380
Involvement	15.78	4.48	15.78	3.60	1.000
Realism	13.55	3.58	12.84	3.62	0.076
Total	56.96	10.05	55.45	8.70	0.192

5 RESULTS

We first present results for each of the self-report measures. In each case we analysed the quantitative results using SPSS v26, to determine statistical significance at the $\alpha = 0.05$ level.

5.1 Presence

We computed values for each subscale of the IPQ, and the overall total, for each participant, in both conditions. We then analysed results for each subscale, and the total score, independently, using paired sample t-tests. The Shapiro-Wilk test for normality was met in each case. We used box plots to identified outliers outside the $Q1 - 1.5 \times IQR$ to $Q3 + 1.5 \times IQR$ range. We retained all identified outliers in each subscale, as we considered that they were a natural part of the data, rather than errors.

Summary statistics for each subscale and total score, for both conditions, are shown in Table 1. No statistically significant difference was found for the *spatial presence* ($t(57) = 0.886$, $p = 0.380$), *involvement* ($t(57) = 0.000$, $p = 1.000$), nor *realism* ($t(57) = 1.807$, $p = 0.076$) subscales. Unsurprisingly, no statistically significant difference was observed for the total presence score ($t(57) = 1.319$, $p = 0.192$).

Although no difference in perceived realism was reported in the IPQ, a number of participants did comment in the interviews on the realism of the toolbox interface, highlighting its physicality. For example, one said “*I preferred using the toolbox than the menu because it was more hands on, and it felt more real as such because you were literally picking out all of the things.*”, and another “*it felt more realistic, almost if you were a real forensic scientist doing the job*”. Another participant noted that the radial menu was “*more like I was playing a game than, like, it was in real life*”. As a counterpoint, some participants also remarked that the lack of weight of the toolbox might undermine its believability: for example, “*I kept expecting there to be weight, which was weird every time I picked it up, but it still felt more real*”.

5.2 NASA Task Load Index

Summary statistics for all subscales, across both conditions, are shown in Table 2, and visualised in Figure 6. The Shapiro-Wilk test indicated that pair-wise differences were normal only for the

Table 2: NASA TLX Summary Statistics

Subscale	Toolbox		Radial		p-value
	Mean	SD	Mean	SD	
Mental Demand	43.97	23.60	37.59	23.42	0.004
Physical Demand	29.91	21.75	19.14	16.20	<0.001
Temporal Demand	28.10	19.28	23.36	17.00	0.026
Performance	27.59	20.57	22.59	19.76	0.052
Effort	46.90	20.41	33.88	20.84	<0.001
Frustration	26.03	22.80	16.38	19.95	0.001

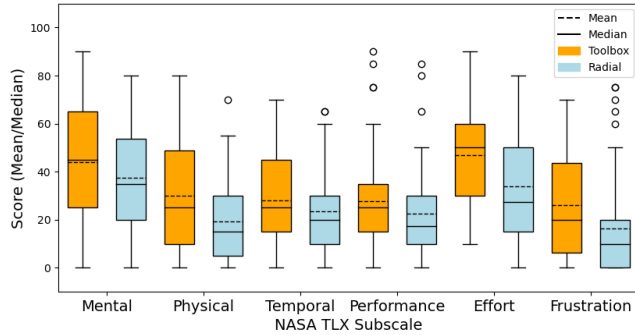


Figure 6: Boxplots for NASA-TLX subscales, for both conditions, showing means and medians

temporal and *effort* subscales, which we analysed using paired t-tests, again retaining outliers. This indicated a statistically significant differences for both the *temporal* subscale ($t(57)=2.290$, $p = 0.026$), and the *effort* subscale ($t(57) = 4.482$, $p<0.001$), where mean values for the toolbox condition were higher in both cases.

The remaining subscales were analysed using Wilcoxon signed rank tests. Statistically significant results were found for *Mental Demand* ($Z=-2.901$, $p=0.004$), *Physical Demand* ($Z=-4.426$, $p<0.001$), and *Frustration* ($Z= -3.43$, $p=0.001$), where scores for the toolbox condition were higher in all cases. The results for the *Performance* subscale were not statistically significant ($Z=-1.940$, $p=0.052$). These results supports our hypothesis H2 for five of the six subscales.

Some comments during interviews supported these findings, and noted a perceived increase in effort or awkwardness when using the toolbox, for example, due to its size, manipulations, or interactions with the environment. One participant stated that it was “*more awkward trying to pick up tools like the camera*” when using the toolbox, and another remarked on the trade off between effort and perceived realism: “*although it was harder work using the toolbox, it made it more real*”. A number of participants noted that placing the toolbox was sometimes awkward, and one suggested “*maybe if there was like a space you could put it on, close to the areas that you had to go to... that sort of central location that you could sort of put your toolbox on*”.

5.3 Task Timing Data

The time taken for participants to complete each task, in each condition, was calculated from the recorded log files, as follows. The duration of the fingerprinting task was defined as the difference between the first teleport performed by the player (which we took to define the start of the tasks), and the last operation of the fingerprint task, as described in the training video (taking a photograph of the fingerprint). The duration of the gel mould task was calculated as the time difference between the end of the fingerprint task and the last operation of the gel mould task (taking a photograph of the footprint), and the duration of the blood swab task was calculated as the time difference between the end of the gel mould task and the

Table 3: Task Timing Summary Statistics

Task	Toolbox		Radial		p-value
	Mean	SD	Mean	SD	
Fingerprint	80.13	27.25	44.39	12.75	<0.001
Footprint	40.97	18.53	28.97	11.98	<0.001
Swab	56.64	15.47	36.88	10.91	<0.001

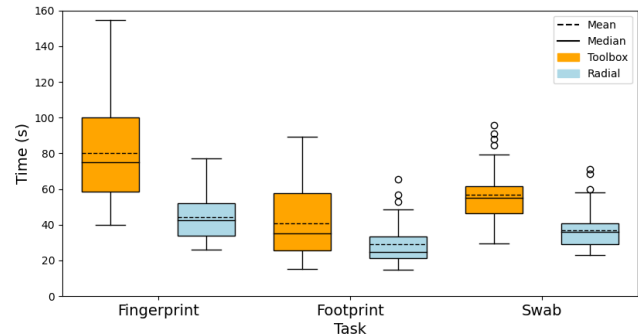


Figure 7: Boxplots for Task Completion Times

last operation of the blood swab (taking a photograph of the blood sample). We identified some instances where participants did not fully complete of all the tasks: we therefore removed timing data for participants who did not, as a minimum, complete each operation of each task (as described in the training video). Fourteen such participants were removed.

Summary statistics for timing data are shown in Table 3, and visualised in Figure 7. The Shapiro-Wilk test indicated that pairwise differences were normal only for the footprint task, which we analysed using a paired t-test, retaining outliers. The t-test indicated that the higher time taken to complete the footprint task in the toolbox condition was statistically significant ($t(43)=4.464$, $p<0.001$). Timings for the remaining tasks were analysed using Wilcoxon signed rank tests. Statistically significant higher times were found for the toolbox condition for both the fingerprint task ($Z= -5.707$, $p<0.001$), and the blood swab task ($Z= -5.765$, $p<0.001$). A number of participants commented on the increased ease and speed of task completion using the radial menu. For example, “*I felt like I had better control over everything, like it was more instant, so I felt like the task went by a lot quicker*”.

6 DISCUSSION

Our results did not show a difference in participants’ sense of presence between the toolbox and radial menu conditions. However, statistically significantly higher scores were recorded for the toolbox condition on all dimensions of the NASA TLX apart from *performance*, and participants also spent longer on each task when using the toolbox. We summarised existing work in diegetic interface design in Section 2.3 (e.g. [11, 22, 30]): our results contrast with those reported by the majority of those studies, which typically saw higher levels or player immersion or presence using diegetic elements. We therefore begin our discussion by considering this disparity.

We firstly note that most existing work on diegetic interface design has considered games, played on 2D screens, where the diegetic world is represented in a projected 3D world space, and non-diegetic elements are presented using screen overlays. However, in VR systems, spatial menus which occupy the same 3D coordinate space as the diegetic world, are more common. This is an analogous interface design, but significantly different from 2D systems, and may partially account for our results. For example, users may perceive that diegetic and non-diegetic components are more distinct on 2D displays as they occupy disparate coordinate spaces. With this in

mind we also draw attention to the work by Salomoni et al. [29,30] which also reported an increased sense of presence for players of a VR shooter game, when using diegetic interface elements. Whilst this study was conducted in VR, at least some of the non-diegetic elements were more akin to 2D screen overlays than spatial menus, and this may also account for differences between their findings and ours: their participants may have perceived a stronger contrast between diegetic and non-diegetic conditions. We note, however, that screen overlays are not common in VR applications, and typically discouraged by developer guidelines [8]. Thus we consider that our comparison is more representative of the interface choices faced by most VR designers.

6.1 Trade offs in Interface Design

Another significant difference between existing works (e.g. Salomoni et al. [30]) and ours is that our diegetic interface design emphasises interaction fidelity. As discussed in Section 2.4, this is considered an important aspect of training applications. Games which use naturalistic movement interfaces have been shown to increase players sense of presence [16,33]; however, we also point to the work by Rogers et al. [28] which indicates that while users enjoy the physicality of movement-based interfaces, more extreme or complex movements may be burdensome or challenging for some.

There is evidence that this may also be the case in our study: participants recorded higher scores on most dimensions of the NASA TLX, which suggests a general perception of elevated workload associated with the toolbox. Participants also spent longer completing tasks, and this may again reflect additional workload. Furthermore, comments by some participants suggest that this could be partly because they spent more time thinking about how to place and manipulate the tools and toolbox, when completing the tasks, (e.g. where best to place the toolbox before opening it). We note that some tools, such as the camera, sometimes required more thoughtful manipulation (for example if they needed to be held at a particular orientation): such considerations could be addressed by refining the design of toolbox (e.g. by adjusting how the tools are positioned therein). We therefore conclude that designing diegetic interfaces in VR applications involves careful contextualised consideration of trade-offs between positive outcomes (such as user engagement) and negatives (such as frustration).

We also note that some participants may have experienced an “uncanny valley” effect when using the toolbox, such as that reported by Bhargava et al. [2]. For example, a number of participants mentioned that the lack of weight of the toolbox felt strange. Such effects are less likely to occur with non-diegetic interface elements, such as the radial menu, which do not imitate familiar real-world objects. Together, such trade-offs and effects may act to limit the gains in presence that can be achieved with higher fidelity diegetic interfaces.

6.2 Implications for Training Applications in VR

Our participants were not CSI students, and so we do not expect that they engaged with the simulation as a serious training experience; however, we assert the validity of our results as being representative of novice users, and draw some implications for the design of VR training applications that involve the frequent selection and manipulation of tools or objects.

In terms of user experience, our results suggest that, for novice users at least, designers should not assume that higher-fidelity diegetic interface elements will necessarily enhance users’ sense of presence or quality of experience. There may be a natural tendency to couple diegetic interface design with high interaction fidelity in training contexts, but this should be carefully considered. This resonates with Alexander et. al’s comments that “*the appropriate level of fidelity for a training system is dependent on the skills or behaviours that are to be trained*” [1], within in the context of VR training. Specifically, results obtained using the NASA TLX, and

for task completion times, seem to indicate lower perceived ease of use with the diegetic condition: it would be difficult to recommend using this interface in practice without a tangible gain in either user experience or learning outcomes. However, it is important to note that this result might partially represent a gap in user knowledge. For example, participants may have found the toolbox more awkward to use because toolbox management is a skill which needs to be explicitly taught to novice crime scene investigators (which we did not include in our pre-condition training). It may then be the case that the radial menu interface is more appropriate for novice users, and the toolbox condition is preferable for more experienced users. However, we also note that while the radial menu system we used was appropriate for the reduced number of tools available in our scenario, the addition of more tools would reduce the size of menu elements, and possibly warrant a more complex menu system, such as a hierarchical design, or larger grid-based layout [21]. Such additions could negatively affect selection speed and/or workloading.

7 LIMITATIONS AND FUTURE WORK

We have highlighted some limitations of our work during our discussion in Section 6: our participants were not selected for experience or knowledge in CSI methods, and are more representative of novice system users. We suggest that CSI students or experienced crime scene investigators might experience the study conditions differently, particularly regarding simulation fidelity, and we propose to examine this in future work.

We also recognise that our results were obtained using a specific training application, and a particular interface structure, and so may not generalise to all training scenarios. However, tool management is a common feature of procedural training simulations, and there is little comparable work on interface design for modern VR systems. We therefore advocate for further research in this area to validate our results in other contexts, and also to explore other diegetic interface elements. Also, the VR experiences which we used in our study were fairly short in duration: we thus propose that evaluations of diegetic interface elements in VR over longer period of use could yield a deeper understanding of user experiences, as well as insights into affects on learning transfer and retention. Finally, we did not directly evaluate the interaction fidelity of our diegetic interface, so comparison with a real-world training exercise using McMahan’s FIFA framework [17] could also be an insightful future study.

8 CONCLUSION

We have presented our study of diegetic interface elements in VR, in which we compared participants’ experiences of a prototype CSI training application using diegetic (toolbox) and non-diegetic (radial menu) tool management and selection mechanisms. Contrary to our primary hypothesis, participants did not report an elevated sense of presence when using the diegetic interface; however, significantly higher levels of workload and completion times were recorded.

Based on these results We have suggested that diegetic interface elements may not necessarily bring the same immediate benefits to users of VR training applications as has been previously reported for games running on 2D displays. This may in part be due to the increased burden associated with the operation of the diegetic interface, which mirrors findings of recent work on high interaction fidelity in VR. We therefore conclude that the design of interfaces for VR training systems could benefit from careful consideration of trade-offs between higher fidelity and negative experiences such as frustration or increased effort, and also that these trade-offs may vary depending on users’ level of skill and knowledge.

There is still little work in the area of diegetic user interfaces, both in VR and 2D applications, and we therefore advocate for more research to support the development of robust and actionable design guidelines in future.

REFERENCES

- [1] A. Alexander, E. Bruny, J. Sidman, and S. Andweil. From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (IITSEC'05)*, 2005.
- [2] A. Bhargava, J. W. Bertrand, A. K. Gramopadhye, K. C. Madathil, and S. V. Babu. Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1418–1427, April 2018. doi: 10.1109/TVCG.2018.2794639
- [3] M. Birk and R. L. Mandryk. Control your game-self: Effects of controller type on enjoyment, motivation, and personality in game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 685–694. ACM, New York, NY, USA, 2013. doi: 10.1145/2470654.2470752
- [4] F. Buttussi and L. Chittaro. Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Transactions on Visualization and Computer Graphics*, 24(2):1063–1076, Feb 2018. doi: 10.1109/TVCG.2017.2653117
- [5] G. Caggianese, L. Gallo, and P. Neroni. Exploring the feasibility of diegetic user interfaces in immersive virtual exhibitions within the cultural heritage. In *2018 14th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pp. 625–631, Nov 2018. doi: 10.1109/SITIS.2018.00101
- [6] T. I. Chowdhury, S. M. S. Ferdous, and J. Quarles. Information recall in a virtual reality disability simulation. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, VRST '17, pp. 37:1–37:10. ACM, New York, NY, USA, 2017. doi: 10.1145/3139131.3139143
- [7] R. Dachselt and A. Hubner. Three-dimensional menus: A survey and taxonomy. *Computers & Graphics*, 31(1):53–65, 2007. doi: 10.1016/j.cag.2006.09.006
- [8] Facebook Technologies. Vrc.mobile.functional.10. Accessed: 2019-11-16.
- [9] A. R. Galloway. *Gaming: Essays on algorithmic culture*. University of Minnesota Press, 2006.
- [10] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909
- [11] I. Iacovides, A. Cox, R. Kennedy, P. Cairns, and C. Jennett. Removing the hud: The impact of non-diegetic game elements and expertise on player involvement. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, pp. 13–22. ACM, 2015. doi: 10.1145/2793107.2793120
- [12] D. Ignacio. Crafting destruction: The evolution of the dead space user interface, 2013.
- [13] L. Jensen and F. Konradsen. A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, 23(4):1515–1529, Jul 2018. doi: 10.1007/s10639-017-9676-0
- [14] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993. doi: 10.1207/s15327108ijap0303
- [15] C. Li, W. Liang, C. Quigley, Y. Zhao, and L. Yu. Earthquake safety training through virtual drills. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1275–1284, April 2017. doi: 10.1109/TVCG.2017.2656958
- [16] R. McGloin and M. Krcmar. The impact of controller naturalness on spatial presence, gamer enjoyment, and perceived realism in a tennis simulation video game. *Presence*, 20(4):309–324, Aug 2011. doi: 10.1162/PRES.a.00053
- [17] R. McMahan. *Exploring the effects of higher-fidelity display and interaction for virtual reality games*. PhD thesis, Comput. Sci., Virginia Tech, Blacksburg, VA, USA, 2011.
- [18] R. P. McMahan, A. J. D. Alon, S. Lazem, R. J. Beaton, D. Machaj, M. Schaefer, M. G. Silva, A. Leal, R. Hagan, and D. A. Bowman. Evaluating natural interaction techniques in video games. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 11–14, March 2010. doi: 10.1109/3DUI.2010.5444727
- [19] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady. Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):626–633, April 2012. doi: 10.1109/TVCG.2012.43
- [20] T. A. Mikropoulos. Presence: a unique characteristic in educational virtual environments. *Virtual Reality*, 10(3):197–206, Dec 2006. doi: 10.1007/s10055-006-0039-1
- [21] P. Monteiro, H. Coelho, G. Gonçalves, M. Melo, and M. Bessa. Comparison of radial and panel menus in virtual reality. *IEEE Access*, 7, 2019. doi: 10.1109/ACCESS.2019.2933055
- [22] M. Peacocke, R. J. Teather, J. Carette, and I. S. MacKenzie. Evaluating the effectiveness of huds and diegetic ammo displays in first-person shooter games. In *2015 IEEE Games Entertainment Media Conference (GEM)*, pp. 1–8, Oct 2015. doi: 10.1109/GEM.2015.7377211
- [23] M. Peacocke, R. J. Teather, J. Carette, I. S. MacKenzie, and V. McArthur. An empirical comparison of first-person shooter information displays: Huds, diegetic displays, and spatial representations. *Entertainment Computing*, 26:41 – 58, 2018. doi: 10.1016/j.entcom.2018.01.003
- [24] L. Pfister and S. Ghellal. Exploring the influence of non-diegetic and diegetic elements on the immersion of 2d games. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, OzCHI '18, pp. 490–494. ACM, 2018. doi: 10.1145/3292147.3292190
- [25] E. D. Ragan, A. Sowndararajan, R. Kopper, and D. A. Bowman. 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, 2010.
- [26] H. Regenbrecht and T. Schubert. Real and illusory interactions enhance presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 11(4):425–434, 2002. doi: 10.1162/105474602760204318
- [27] A. Ricca, A. Chellali, and S. Otmane. Study of interaction fidelity for two viewpoint changing techniques in a virtual biopsy trainer. In *2017 IEEE Virtual Reality (VR)*, pp. 227–228, March 2017.
- [28] K. Rogers, J. Funke, J. Frommel, S. Stamm, and M. Weber. 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*, CHI '19, pp. 414:1–414:14. ACM, New York, NY, USA, 2019. doi: 10.1145/3290605.3300644
- [29] P. Salomoni, C. Prandi, M. Rocchetti, L. Casanova, and L. Marchetti. Assessing the efficacy of a diegetic game interface with oculus rift. In *2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC)*, pp. 387–392, Jan 2016. doi: 10.1109/CCNC.2016.7444811
- [30] P. Salomoni, C. Prandi, M. Rocchetti, L. Casanova, L. Marchetti, and G. Marfia. Diegetic user interfaces for virtual environments with hmds: a user experience study with oculus rift. *Journal on Multimodal User Interfaces*, 11(2):173–184, Jun 2017. doi: 10.1007/s12193-016-0236-5
- [31] A. Santos, T. Zarraonandia, P. Díaz, and I. Aedo. A comparative study of menus in virtual reality environments. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, pp. 294–299. ACM, New York, NY, USA, 2017. doi: 10.1145/3132272.3132277
- [32] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, 2001. doi: 10.1162/105474601300343603
- [33] D. M. Shafer, C. P. Carbonara, and L. Popova. Spatial presence and perceived reality as predictors of motion-based video game enjoyment. *Presence: Teleoperators and Virtual Environments*, 20(6):591–619, 2011.
- [34] R. Skarbez, F. P. Brooks, Jr., and M. C. Whitton. A survey of presence and related concepts. *ACM Comput. Surv.*, 50(6):96:1–96:39, Nov. 2017. doi: 10.1145/3134301
- [35] M. Slater, P. Khanna, J. Mortensen, and I. Yu. Visual realism enhances realistic response in an immersive virtual environment. *IEEE Computer Graphics and Applications*, 29(3):76–84, May 2009. doi: 10.1109/MCG.2009.55

- [36] M. Slater and S. Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6(6):603–616, 1997. doi: 10.1162/pres.1997.6.6.603
- [37] M. Speicher, P. Hell, F. Daiber, A. Simeone, and A. Krüger. A virtual reality shopping experience using the apartment metaphor. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces, AVI '18*, pp. 17:1–17:9. ACM, 2018. doi: 10.1145/3206505.3206518
- [38] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4):73–93, 1992. doi: 10.1111/j.1460-2466.1992.tb00812.x
- [39] J. Stevens and J. Kincaid. The relationship between presence and performance in virtual simulation training. *Open Journal of Modelling and Simulation*, pp. 41–48, 2015. doi: 10.4236/ojmsi.2015.32005
- [40] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks, Jr. Walking > walking-in-place > flying, in virtual environments. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '99*, pp. 359–364. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1999. doi: 10.1145/311535.311589