2018-11-27

Using mixed reality displays for observational learning of motor skills: A design research approach enhancing memory recall and usability

Watson, PWL

http://hdl.handle.net/10026.1/12950

Using mixed reality displays for observational learning of motor skills. A design research approach enhancing memory recall and usability.

Abstract

When learning an action sequence, observing a demonstration informs the knowledge of movement execution and enhances the efficiency of motor skill acquisition. 3D virtual learning environments offer more opportunities for motor skill training as they afford observational learning. Mixed reality platforms (virtual reality, desktop PC etc.) that render 3D virtual environments can therefore increase accessibility of observational content. To explore effectiveness of these platforms to facilitate observational learning of action sequences, we developed the Recovery Position Application[1] (RPA) at the Interactive System Studio, University of Plymouth. The RPA was originally designed for mobile virtual reality. The RPA displays two virtual avatars performing the steps of the recovery position. We present the design of content and interaction informed by research into observational learning of motor skills. To formatively evaluate the current functional prototype, and potential use within an educational context, RPA was tested on three different platforms. Mobile VR (N=20), desktop PC (N=20) and video recording (N=21). Memory recall of movements were recorded and the usability of the RPA was investigated. Across all three platforms, the average recall of demonstrated information was 61.88%, after using the application for 10 minutes. No significant differences between recall rate was identified between platforms. Participant's responses were positive or very positive for both application effectiveness as a learning resource and for ease of use. These results are

discussed in regards to the future development of the RPA and guidelines for virtual demonstration content.

Keywords

Introduction

In training and education, there are many instances where students will need to imitate a performance from a demonstrator. For example, to gain understanding of how to use laboratory equipment, use computer software, or to acquire a set of motor skills for sport etc. Demonstration-based training (DBT) (Rosen et al., 2010) requires effective delivery of observational content for students to learn from.. A *demonstration* is a "dynamic example of partial – or whole-task performance" that conveys the required knowledge, skills, and attitudes to the learner. The two learning opportunities are when the student observes the demonstration and when any activity supplements the understanding of this performance either pre, during, or post demonstration.

DBT is a common approach used to teach motor skills. For example, in acquiring a set of dance movements, the teacher will demonstrate an action and then ask the student to imitate said action for practice. Central to this process is the use of observational learning by the student. Although physically practicing a motor sequence grants an implicit long-term memory of movements (Boutin et al., 2010; Wulf & Schmidt, 1997), the addition of

_____

observation enhances the efficiency of motor skill learning (Ashford, Bennett & Davids, 2006).

Through observation, an individual can acquire a mental representation of a motor skill to cue imitation (Sheffield, 1961) and correct errors. Fitts' and Posner's (1967) model for learning motor skills describes three typical stages: cognitive, associative, and automatic. This model establishes that cognitive representation is important at the beginning of motor skill development where knowledge of movement positions and goals are limited. Later into motor skill development the learner may still benefit from more demonstrations as they refine their movements, but application of technique and feedback become more critical to learning. Knowledge of how to execute an action does not mean an individual is proficient at performing said action. A student will inform their own progress throughout training with own self-analysis and feedback from instructors. Through sleep, cognitive representation and motor neuron information from physical practice will consolidate (Walker et al., 2002). Students will therefore normally need to practice over many days and weeks to develop a motor skill. Applying a motor skill to a variety of scenarios will develop generalisability of use. Identifying when a student needs to vary their training, or focus on a specific detail, is informed by the goals of the student and the judgement of the instructor (Williams & Hodges, 2005). Demonstrations aid the process of motor skill development by providing a mental representation of actions to inform movement goals during practice sessions and information to help define criteria for feedback. Therefore, the timing and content of a

demonstration will depend on the structure of a training program and the current level of experience the student with a motor skill.

If cognitive representation is an outcome from observing a demonstration, the representation will be encoded into memory (Fitts & Posner, 1967; Bandura, 1977). This memory may be symbolic and subject to decay (if not practiced or rehearsed), but will provide information for the user to decode, interpret and subsequently imitate, by providing familiarity and valuable analysis not available while performing the actions (Bandura, 1977; Elliott et al., 2011).

3D virtual environments with animated avatars extend the opportunities for observing demonstrated content inside and outside the classroom, affording realistic spatial knowledge representations by replicating real world perspective and lighting (Dalgarno & Lee, 2010). Desktop PCs and mobile devices, like tablets and smartphones, provide a variety of platforms to present virtual demonstrations. Immersive technologies like head mounted displays (HMD) can visually and audibly envelope the user as if they were present within an actual environment, mirroring a real-world viewpoint. This viewpoint is egocentric (displays objects in relation to the user) which will aid general mapping of environments, including allocentric representations (objects in relation to each other but not to self) (Epstein et al., 2017). Yearly advancements in smartphone performance has increased accessibility to devices that can render 3D virtual environments. The improved functionality has enabled smartphone focused virtual reality (VR) platforms like Google Cardboard (Google Cardboard, 2018) that has been widely adopted for entertainment. Mobile VR has significant potential to aid observational learning by providing virtual demonstrations inside and outside the classroom. This form of demonstration supports in class

4

training, facilitates DBT at long distance and enables observational learning for independent study.

To investigate the use of virtual content as a tool for observational learning of motor skills, the Recovery Position Application (RPA) was developed for the Google Cardboard platform. The aim of this application is to show the recovery position action sequence, for the user to observe and memorise. Within the framework of DBT, this application has one function: display the demonstration. To apply relevant theory to the design of the RPA, a software development approach was used. The design process of software development establishes the requirements (needs) of the software based on the business case (problem domain). These requirements are then broken down into features and functions that consider the target audience, software and the hardware of the technology. Once developed, these features are then tested to evaluate their implementation. The key requirement of RPA was to utilise a smartphone as the source for demonstration content. Through research and analysis of both observational learning and the target hardware, the RPA was constructed by an interdisciplinary development team.

When designing any technology to support education and training, the usability of the hardware and software are as important as the content. The usability of a system describes how effectively and efficiently the desired goals can be attained in a specified context of use, and the user perception of this process (ISO 9241-11, 2018). This definition values both objective performance and perceived achievement. Poor usability will deter both the provider and receiver of information in an educational setting. If students cannot easily learn how to use a tool, then the instructor will need to spend more time educating and troubleshooting the tool (Akçayır & Akçayır, 2017). Outside of a structured lesson,

students may avoid the technology entirely as they do not have the technical support of the instructor. In the context of an immersive virtual environment, technical frustrations or usability issues can break the psychological sense of presence, where the user objectively sees and subjectively feels that they are not in the real world, but the synthetic virtual setting (Slater, 2003). Breaking presence will engage the user with the real world and therefore distract them from the virtual experience.

To evaluate the direction of applications in development, regular testing is crucial. Features developed from informed design still need to be tested to see if they are fit for purpose, and usable by the target audience within a given context. Lessons learnt early, can inform future iterations of development. To formatively evaluate the functional prototype RPA and inform the continued design process of the application, a usability study was conducted.

The aims of this study were to:

1. Measure the memory recall of movements observed.
2. Evaluate usability; ease of use, perceived effectiveness of learning.
3. To compare RPA across different platforms.
4. Establish guidelines for developing virtual demonstration-based content and delivery.

Design of the Recovery Position Application

The RPA displays two virtual avatars performing the recovery position sequence. The avatar performing the sequence is named "Helper". The avatar placed into the recovery position is named "Casualty". The design of RPA was based upon the needs of observational learning within a DBT framework and the hardware considerations for mobile VR. To view a video, walking through the application. Please go to: https://goo.gl/F5phhE. The two areas of focus to inform the design process were:

1. Technical limitations of the hardware and Google Carboard platform.
2. Requirements of successful observational learning of a demonstration.

Hardware Considerations of Mobile Virtual Reality

Mixed reality describes a broad set of technologies that combine real and virtual worlds for interactive experiences. The spectrum of platforms ranges from entirely synthetic, virtual environments to completely real environments (Milgram & Kishino, 1994). At one extremity of this spectrum is virtual reality (VR). VR is a computer generated synthetic world that responds realistically to human senses, and thus creates the illusion that the user is in a new reality (Slater, 2014). The term "Immersive display" describes how the user's visual sense is surrounded by the virtual world with replication of a stereoscopic view (a display for each eye) and human perspective (closer objects appear bigger) (Slater & Sanchez-Vives, 2016). An example of an immersive display would be a head mounted display (HMD) or wall projection system like the CAVE (Cave Automatic Virtual Environment) (See.Fig.1). In such displays, the position and rotation of the user's head is tracked in 3D space and updates the display accordingly. This positional and rotational tracking is described as six degrees of

freedom (6-DOF). This relates to the 3 axial points (x, y, z) that are recorded for *both* position and rotation. As the user turns or moves their head, it appears as if they are turning their head within the virtual world.

Fig.1 Virtual Reality Examples: (Left) HTC Vive headset that is tracked in 3D space (Vive.com, 2018). (Right) A CAVE where the displayed image is projected onto a surrounding surface from the perspective of the user (Visbox.com, 2018).
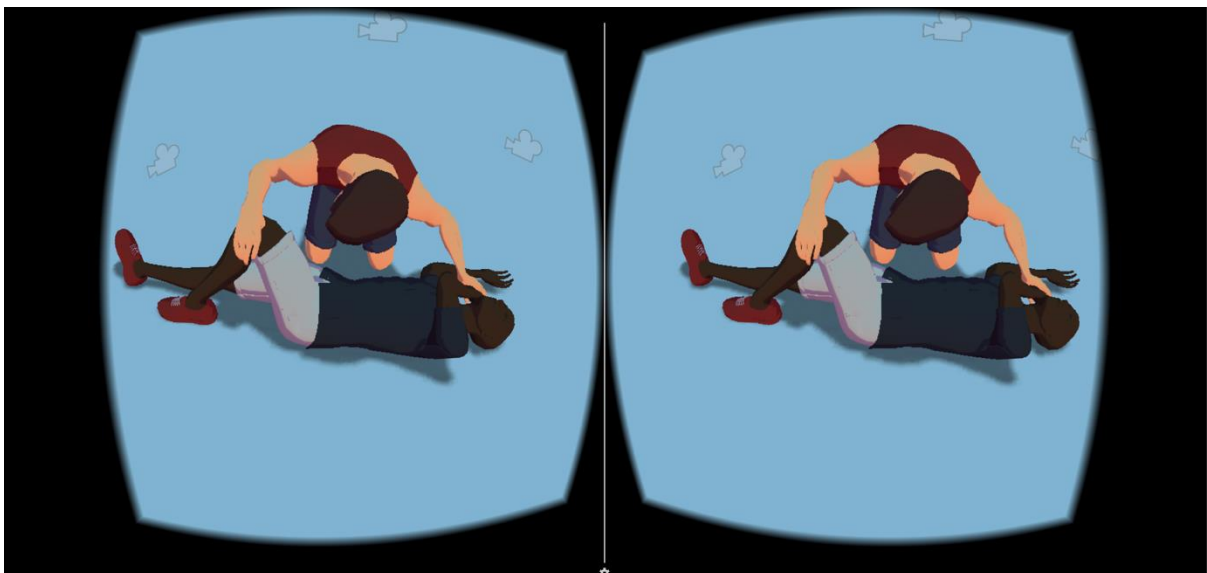

Fig.2. Shows how immersive displays render to each eye on the same display creating stereoscopic vision. Graphics of application are informative but optimised.

A mobile phone can deliver a similar, *mobile VR* experience. A mobile phone is placed inside a headset close to the user's eyes to become the display (See Fig.3). Although a mobile phone display will have much less graphical power than a PC, it will need to render acceptable framerates at clear resolutions to minimise visual lag of movements. Visual lag can cause simulation sickness (Davis, Nesbitt & Nalivaiko, 2014). This does raise a design consideration that graphics should be informative but need to be optimised stylistically to maintain graphical performance (See Fig.2).

Fig 3. (Left) Shows how a mobile phone is placed as the display for Google Cardboard head mounted display. (Right) Shows the placement of the Google Cardboard display in use.

Selection techniques refer to how a user interacts with the graphical user interface of an application to make choices of how to progress and change settings. Although some dedicated mobile VR hardware configurations will have a connected controller or a single HMD button to press, this is not the standard. To increase accessibility of the RPA (Both dependent on technology and potential disability of a student), the assumption will be that users will have no more than a smartphone and an HMD housing. For this reason, a time on target selection method was used. This method allows all selection choices to be controlled with the movement of the user's head. A black circle in the centre of view acts as the user's

reticule (which represents the relative centre of the display). When the user rotates their

head and positions the reticule over selectable icons it will enlarge, disappear and then

slowly draw a circle. Once the circle is complete, the icon is selected. The delay caused by

the circle draw facilitates an intended action and reduces the chance of accidental selection
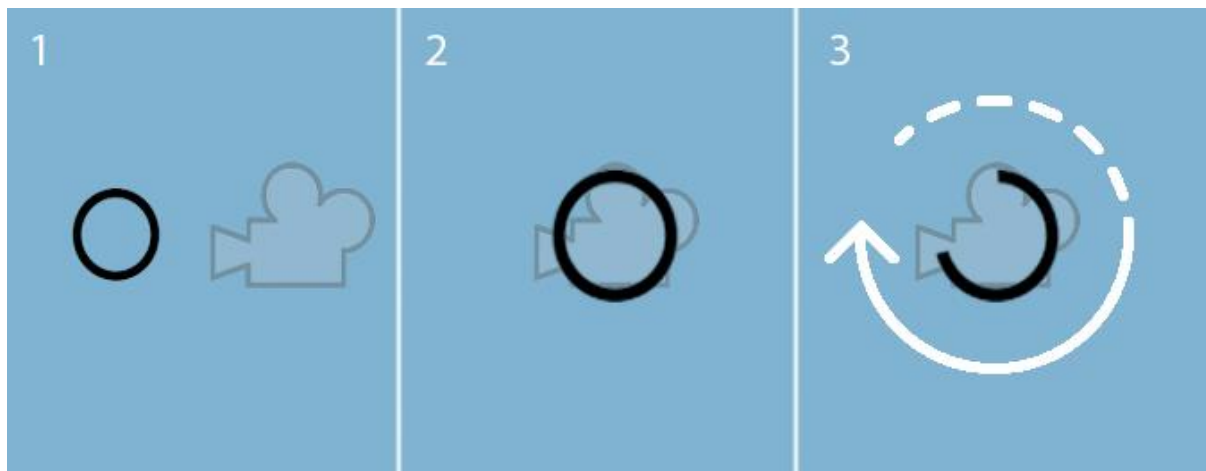
(See Fig.4).

Fig.4. Image sequence showing time on target selection method of camera icon (panel 1).

User reticule is positioned central to the viewport (Black circle, panel 1). User positions

reticule over camera icon (panel 2). Reticule expands to provide feedback that icon is

selectable. The reticule then redraws itself over 2 seconds (panel 3, white arrow shows

direction of redraw). When the circle is fully redrawn, the item is selected.

A comparative limitation of mobile VR is that only rotational information is tracked in

3D space (through the phone's accelerometer), positional information is not. This is known

as three degrees of freedom (3-DOF). In practical terms, the user can rotate their head but

there will be no visual feedback of translational movement (walking, crouching etc).

Without the visual update, translational movement could lead to simulation sickness and

breaking the sense of presence. Due to this limitation, the user should be seated when using

mobile VR. To enable navigation of the 3D virtual environment without a controller input or

tracking of translational movement, a teleportation system was developed. Teleportation locations are visualised by camera icons. Once a camera icon is selected, the user will teleport to that location.

Delivery of Observational Content

This initial iteration of the RPA delivers a virtual demonstration to create a mental representation of the action sequence. To inform design and implementation we used an observational learning approach, key elements and considerations are described below.

Cognitive load theory describes how the processing of the learning task (intrinsic), task presentation (extraneous), and the mental resources devoted to assimilating information into long term memory (germane) utilise an amount of information within working memory (Sweller,1988; Van Merrienboer & Sweller, 2005). As cognitive load increases, fewer mental resources are available to explore learning scenarios and assimilate information. Showing all elements of a concept can be too much new information for a student to interact with and apply (Pollock, Chandler & Sweller, 2002). Practically, this suggests that breaking down a movement sequence into individual actions may reduce the intrinsic cognitive load (Yang et al., 2013). When using technologies or learning materials that are unknown to the user, there will be an increase in extraneous cognitive load. Extraneous cognitive load can also be increased by overloading sensory streams. For example, if all information is presented visually (writing, diagrams, animation) then the visual stream can be overloaded. Directing some information through the auditory stream (written text to verbal oration) reduces extraneous cognitive load. The RPA uses narration to

accompany the visual demonstration that describes as actions are being performed by the avatars.

When observing actions, students perceive the spatial coordinates of an instructor's movement in relation to the demonstrator's body. This provides a reference for body position and speed of actions. To imitate these actions, the student must mentally map this information to their own body. This requires transcoding information from an allocentric spatial frame (objects are located relative to one another) to an egocentric spatial frame (objects are located relative to the learner's body) (Willingham, 1998). The transcoding will require an amount of cognitive load based on the learner's ability to mentally rotate and familiarity with the action (Krause & Kobow, 2013). This increase in cognitive effort is evidenced in motor skill imitation studies. Participants took less time to imitate hand actions when viewed from an egocentric spatial frame, compared to an allocentric one (Jackson, Meltzoff & Decety, 2006). Physically aligning the allocentric spatial frame of the demonstrator with the egocentric spatial frame of the student will reduce the extraneous cognitive load (Krause & Kobow, 2013). Therefore, the observational content within a virtual environment should allow multiple vantage points that enable user navigation between allocentric and egocentric perspectives as in our implementation.

Learner autonomy over navigation around an object (Brooks et al., 1999) has been shown to improve the memory recall of complex 3D objects and spatial layouts. Participants that memorised the layout of a virtual building recalled more when in control of navigation. Similarly, when participants rotated a virtual inner ear model (Jang et al., 2016), those that had control over the direction of rotation were able to draw this anatomical structure more fully. This prior research established that autonomy over navigation and flow of an

experience can aid the spatial and episodic memory of what is observed. Virtual observational content should therefore give control to the learner on how they travel, choice of perspective, and the pace at which they explore the observed information. In our implementation users are free to teleport between observational viewpoints and control over when the next sequence performed is enacted (See Fig.5).

Fig.5. Users can teleport to the camera locations (Marked "C") at any point during the demonstration. Demonstration will not progress to the next movement until users select the "Next Step" icon symbolised by a white circle (Marked "S").

When observing actions for later memory retrieval, recognition or imitation, it is important that movement is demonstrated accurately in terms of body posture and time spent transitioning between poses. From brief observation of a body posture, and with no attempt to imitate, we can accurately remember and recognise action poses (Urgolites & Wood, 2013). Action can be understood even when abstracted into 2D images or when the action is described verbally. However, motor skill acquisition is more effectively taught through animations than still pictures (Höffler & Leutner, 2007). As long as the focus of the

subject matter (in this case analysis of human movement) is represented accurately, then there is little benefit in raising the fidelity of graphics (Norman, Dore & Grierson, 2012).

Neurological studies of observing actions suggest that similar neurons fire when an action is performed and when passively observed (Rizzolatti & Sinigaglia, 2010). Through this *mirror mechanism,* it is suggested that we internally simulate performing an observed action to predict possible action. This neurological representation is mediated by our understanding of the observed action. For example, Lacoboni *et al.* (2005) showed participants images of grasping a mug with no context and within the context of breakfast. FMRI (Functional Magnetic Resonance Imaging) scans showed a stronger activation of mirror neurons when the context was established. The observer's own goals can influence the pattern of mirror neurons that discharge. In previous work (Molenberghs et al., 2012), participants were asked to observe hand actions under three different mindsets: Understand the meaning of an action, observe physical features, or passively view the actions. FMRI scans showed subtle variations in mirror neuron discharge dependant on this mindset. These two studies show that higher level cognitive process (mindset and context of a situation) mediate the neurological representation of movement when observing actions. If participants are told that they are playing a computer game to improve their golf putting ability as opposed to simply enjoying the experience, they will show better real-world improvement (Fery & Ponserre, 2001). Establishing the context of the learning activity and observed demonstration motivates the user to learn as well as aid the cognitive representation of actions. In the RPA, context is established at the outset by combining an instructional voice over and text overlay that informs the user of application's educational purpose and goals.

To summarise, immersive displays facilitate an egocentric perspective within a 3D virtual environment and so replicate observations of movement as if in the real world. The mental representations of these movements may aid understanding of actions throughout motor skill development. The core requirements of effective observational learning of action sequences via demonstration are: breaking down a movement sequence into manageable chunks, spatial representation of actions, manipulation of spatial frames of reference, user control, accurate representation and context.

Methodology

Procedure

To explore the effectiveness and usability of the RPA as a tool for observational learning, three conditions were used in a between groups design. Group one (N=20, 2 female), titled "Mobile VR", used mobile VR to interact with the application. Group two (N=20, 1 female), titled "Non-Immersive", used a desktop PC display with a mouse and Keyboard for navigation. Group 3 (N=21, 3 female), titled "Video", watched a video recording of the RPA on a desktop PC display. The use of the Non-Immersive group was to test the difference in recall between an immersive mobile VR and a non-immersive display. The video group was used to explore the role of autonomy on memory recall compared to the desktop PC group. By watching a video of the RPA, the participants would be viewing the same content but without control over the flow of information and choice of perspective. The background of all participants were a mixture of computing students and employees of software development companies.

The same procedure was used for all groups: Participants filled in a questionnaire that extracted basic demographic information, including their opinion on current ability to perform the recovery position and perceived knowledge of related technology. Verbal instructions on using the RPA were given. Participants were then seated and asked to play the "Interaction" mode twice. During *interaction* mode, participants selected when to move between the individual steps of the recovery position. Participant could teleport between viewpoints to change their viewing angle and each step of the visual demonstration was accompanied by a verbal description. Participants were then invited to review the "Observation" mode with the avatars enacting the recovery position action steps in

sequence automatically. No verbal description was present in this mode. This process took participants 15 minutes to complete.

Post exposure to the RPA, we tested participant's memory recall of the recovery position. Participants had the choice to write down what they remembered or verbally report it to the investigator. Participants then filled in a questionnaire exploring usability of the application and perceived usefulness as an educational tool. The questionnaire used a five point Likert scale for 11 questions (9 for usability, 2 for perceived usefulness). For the Mobile VR group, the participants were then interviewed on their experience in using the Recovery Position Application. With participants' permission, both memory recall test and interview were digitally recorded.

Data Analysis

In total, the RPA delivers 27 details about the recovery position (See Table 1). To help segment the type of information participants successfully recalled, these details were divided into four categories: Movement, Assessment, Instructional, and Supportive. Movement details relate to specific visual or audible instructions for the helper avatar to position the casualty avatar. The Assessment details relate to judgements within the demonstrated scenario. For example, the participant may be instructed "Only proceed if the casualty is breathing normally". Instructional details relate to tasks that do not explicitly state how they should be physically performed. For example, "Call an ambulance". Supportive details are extra information that explain why a movement is being performed.

Information Delivered Through the Application

M = Movement, A = Assessment, I = Instructional, S = Supportive

Step No.	Detail No.	Detail Description	Info Type
01	1	Check the area poses no risk to yourself	A
	2	Check that the casualty is breathing	A
	3	Gently tilting the head back	M
	4	Listen and feel for breath on your cheek	M
	5	Look for movement in the chest	I
	6	Only proceed if they are breathing normally	A
02	7	Select arm closest to you	M
	8	Place at right angle to casualty's body	M
	9	Palm facing up	M
03	10	Select hand furthest from you	M
	11	Bring across casualty's body	M
	12	Place back of casualty's hand against patient's cheek	M
04	13	Grab knee furthest from you	M
	14	Raise it up	M
	15	Until foot is flat on the floor	M
05	16	Gently roll casualty towards you	M
	17	By pulling on the knee	M
	18	Support the casualty's head with your hand during this manoeuvre	M
06	19	Tilt the head	M
	20	By lifting the chin	M
	21	Ensure airway is open	A
	22	Check for normal breathing	A
07	23	Select top leg	M
	24	Bring out at right angle	M
	25	To support the casualty	S
08	26	Call an ambulance	I
	27	Monitor the casualty	A

Table 1. Table to show information delivered through the Recovery Position Application. The "Step No" of Table 1 categorises a group of details delivered in one section.

Results

Information Recall

Memory Recall % for Information Types

Information Type	Mobile VR	Non-Immersive	Video	Average across conditions (Mean)
Movement Recall %	73.61	63.61	71.43	69.55
Assessment Recall %	47.50	43.33	59.52	50.12
Instructional Recall %	45.00	35.00	50.00	43.33
Supportive Recall %	30.00	50.00	67.67	48.89
Overall Recall % (Mean)	62.14	56.48	67.02	61.88

Table 2. Comparison between groups of recalled information categorised by information type.

Table 2 shows the overall recollection rates of movement information as 69.55 %. This is 7.67% higher than the overall recall rate of all the questions (61.88%). Assessment (50.12%), instructional (43.33%), and supportive information (48.89%) were recalled less. A one-way between groups ANOVA was used to compared memory recall across the three groups. The difference in recall was not statistically significant (F=2.64, P= 0.079). In all groups, *movement* details were more frequently recollected than others.
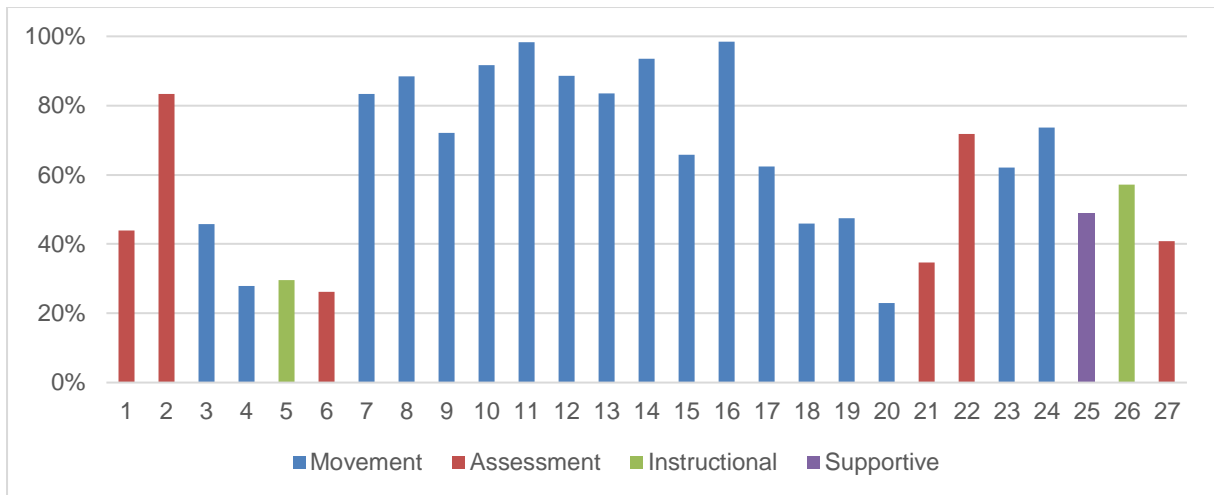
Fig.6. Average recall (across all groups) of details presented in the Recovery Position Application. These details are colour coded to the information type
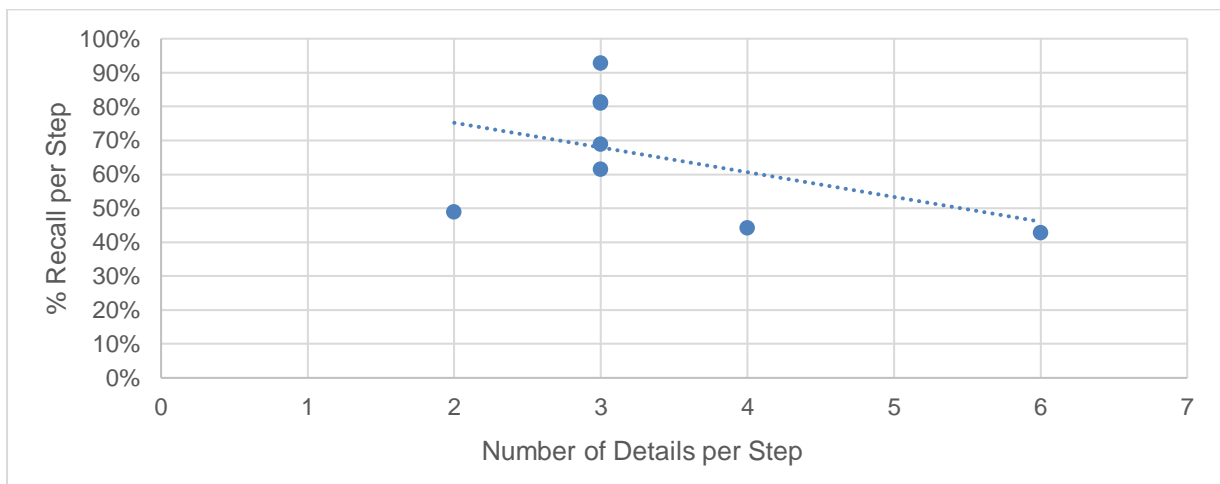

Fig.7. As the amount of details in each step increased, the % of information recalled decreased.

Fig.6 shows that two of the movement details were recalled comparatively poorly. These were step 01, detail no. 4 (27.94%) and step 6, detail no.20 (22.94%). These details were all associated with subtle positioning of the casualty's head. These details were also delivered in steps 1 and 6, which contained more information than the other steps. Fig.7 shows that there is a moderate negative correlation (-0.46) when comparing the amount of information shown per step and the memory recall of that information.

Usability

To explore participant's user experience of the Recovery Position application, the system usability scale survey (Brooke, 1996) was used. Each question was answered with a five point Likert scale. For each question, the modal average of this scoring was calculated.
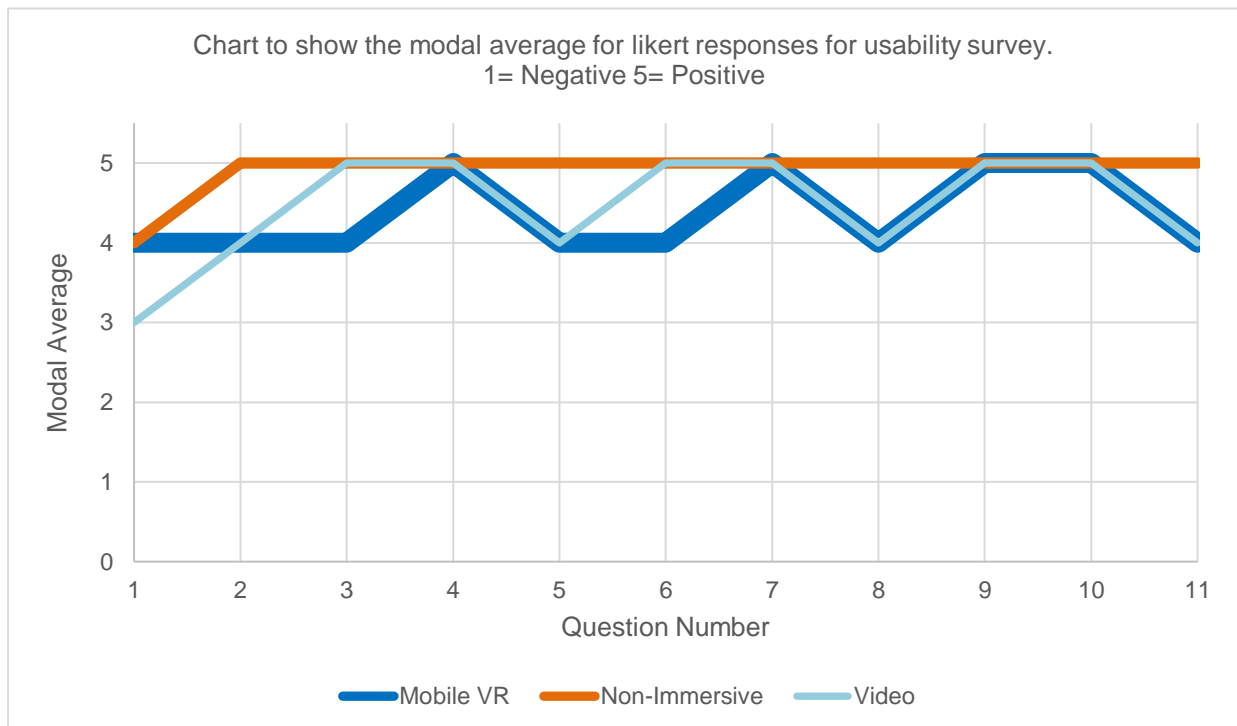

Fig.8. Results of Likert questionnaire detailing participants' perception of the Recovery Position application for all three groups. This indicates a general positive perception to the usability of the application independent of the platform.

For all groups, the usability questionnaire showed positive attitudes towards the use of the RPA for education (See Fig.8). Questions 2 – 10 examine the RPA's ease of use. Modal averages were 4 or 5, representing positive and very positive. This indicates participants found the application easy to use and felt confident in utilising the functions. Questions 1 and 11 relate to the participant's perception on whether they would use the application frequently and recommend it to others. Participant's modal average responses for the

mobile VR and Non-Immersive groups were 4 and 5 respectively. This indicates that participants would use the RPA frequently and recommend it to others in a mobile VR or desktop format. Question 1 of the usability survey for the Video group had a modal average of 3. This shows that for the video format, participants are neutral in their option to use the RPA frequently and may not use a video format at frequently as a desktop or mobile VR platform.

Autonomy

Group three (video) were asked a further four questions, in the post exposure questionnaire, to help evaluate usability when not having control over navigation and flow of content.

- Q1 - Would you prefer control of navigation and flow of information?
- Q2 – Why is this?
- Q3 - Did you find the lack of control in navigation and flow of information frustrating?
- Q4 – Why is this?

85.7% responded "yes" to Q1. 42.9% of participants responded "yes" to Q3. This suggest that participants would prefer control of the flow of information, but 57.1% did not find the lack of control frustrating. In response to Q2, 52.4% participants valued: "learning at my own pace", 19% valued being able to repeat a step when needed, and 15.3% valued control over the camera viewpoint. The lack of these elements was cited as the cause for frustration

when answering Q4. However, those that were not frustrated, explained that information was delivered at a suitable pace.

Interviews

To explore the usability of the RPA in more detail, we interviewed group one (mobile VR) on their experience. Below is a synopsis of the key findings:

Notably, 95.0% perceived that they knew significantly more about the recovery position after using the application. 95.0% reported that they found the software easy to use and that the controls did not interfere or distract form the material delivered.

Participants reported that they predominantly only used two view positions. These were the front and back, as they facilitated oversight of the entire action sequence. A common request was to add two more camera positions. One directly above the demonstration for overview, and one directly from the viewpoint of the "Helper" avatar. The latter viewpoint was to experience observation of the action sequence, as if performing the recovery position. Viewpoints positioned near the head of the casualty were deemed too close to observe any meaningful details by some participants. The proximity of these viewpoints to the casualty forced users to translate their head position for better vantage. This highlights a limitation of mobile VR. The positional translation of the HMD is not tracked in 3D space and limits the user's natural head movement to rotations only. Participants described this limitation as frustrating.

Participants viewed the graphics as believable, even though they did not describe the style as realistic. A key driver for this believability was the perceived realism of the

animations. One participant elaborated on how they expected the graphics to not be realistic: "It might be cartoony, but that is what you expect from an app. You don't expect a real-life person".

A highlighted feature for improvement that many participants requested was a repeat step function. In the current version of the RPA, you cannot repeat a single step. Participants might miss a detail and would have to repeat the entire sequence to review a step.

Having both audio description and visual demonstration was perceived as an effective combination for information delivery. Some participants admitted that they only listened to the descriptions as they found this easier to assimilate. Some ignored the audio description completely. The majority found useful information in both. Participants also noted that a reason why they did not mention the first detail (check that the area is safe) is because there was no demonstration of this action, just audio. Secondly, there was no visible danger in the environment.

Discussion

Across all three groups, participants were able to recall most of the correct body positions for each action and the correct sequence order, despite using the RPA for a short duration (10 minutes). This suggests that a cognitive representation of the recovery position was developed effectively through using this application. Other details (assessment, instructional, and supportive) were recalled less frequently. The focus of this study was on movement recall, and the participants were informed of this prior to using RPA. A possible result could be that participants ignored non-movement information due to this instruction. Also, representation of non-movement information was only through audio and not through animated content like the movement sequence. Interviews from participants suggest that some ignored information that was not visually demonstrated by the avatars, and many perceived that a combination of audio and visual representation to be effective at delivering information. Therefore, information delivered only through audio may create weaker memory hooks, or emphasis, than through the combination of visual and audio. Alternatively, there may have been too much information in the steps that provided non-movement details. For example, the memory recall from step 01 had the lowest recall rate and the highest amount of details. This indicates that this step had too much information and superseded the limited working memory (Miller, 1956). When working memory is exceeded the individual will either ignore any extra information or may develop a method to organise it into smaller chunks (Yang et al., 2013). Logically, a reduction in information per step, may improve recall. Additionally, many participants mentioned through interviews that they would like the functionality to repeat an individual step. This function could help users to retain more details by enabling more exposure to information when needed.

25

There was no significant difference in recall between the three groups. Each group contained a similar sample (age range, background, gender etc). Any difference in recall between the groups could be explained through variances within each group (prior knowledge of the recovery position, exposure to technology, ability to mentally rotate 3D structures etc.). The similar recall across the three groups suggests that the RPA could be used on multiple mixed reality platforms with similar effect on demonstration recall. Within the DBT framework, demonstrations of action sequences could be supported by many display devices, depending on the needs of the training, or the resources at hand. Of interest, is that a more immersive device (mobile VR) did not aid recall of action sequences compared to a non-immersive display (desktop PC display). This may suggest that a PC display provides enough spatial information and effective egocentric/allocentric perspectives to create a mental representation of observed movements. In addition, the limitations of mobile VR (translational movements not tracked, small field of view etc.) may weaken any benefits a dedicated immersive VR (HTC Vive, CAVE) set up may facilitate for observational learning of demonstrations.

The video condition removed participant autonomy over navigation and flow of information. Survey results and interviews suggested frequently that users would like more autonomy when exploring the RPA. However, removing autonomy did not translate to poorer memory recall. A more complicated movement set or longer exposure to the RPA in this format may cause a negative effect on memory recall due to a lack of autonomy. However, for short (10 minute) demonstrations, this study suggests that autonomy of information has little impact on memory recall and a minor negative effect on usability.

When observing the RPA demonstration, actions performing head adjustments to the casualty were least recalled. In step 01, the higher frequency of details (06) could have reduced overall recall. However, head movement details were also poorly recalled in step 06 which had four details. This can be explained in terms of the learner's goals. In *goal-directed imitation* theory (GOADI) (Wohlschläger, Gattis & Bekkering, 2003), the goal and intent of the movement supersedes the act of imitation. For example, in flicking a light switch on, the individual is not concerned with how this is achieved (correct arm direction, which part of the hand to use etc.) but focused on the goal of the action (move switch up). Similarly, in this study, participants are acquiring a mental representation of moving a body into the recovery position. The smaller details of this act, although important, may not be the focus of the participant's goal. In this case placing a casualty into the end body position of the recovery position is the user's primary goal. Feedback form interviews also suggests that the movement of the helper avatar's hands were not clear, when viewed from an obstructed angle. Separating head and hand actions into a single step, with more detail, could aid recall of these movements.

User feedback recorded through post use surveys showed a positive perception for the RPA in the context of usability and an educational tool. This was across all groups using immersive and non-immersive displays. Participants did receive training in how to use the time on target selection method through a brief description. The duration of basic training was minimal and took no longer than 3 minutes. This suggests that the RPA does not impose significant additional barriers to the learning of the content. The technology platform, design of RPA, and introductory tutorial may have minimised the extraneous cognitive load. The ease of use and positive perception across mobile VR and desktop PC displays indicates that this type of educational tool could be an effective resource outside of the classroom.

This type of tool could therefore be useful as a revision tool for DBT lessons, and a primer before movement is enacted. It could also aid other teaching frameworks that require pre-session study like the flipped classroom.

Conclusion

By analysing the interaction design requirements and hardware limitations, we have created an application that shows good usability at an early development stage. By translating the principles of effective observational learning into application features, the RPA demonstrates the effectiveness of mixed reality devices (both immersive and non-immersive) to deliver virtual demonstrations of action sequences for observational learning. The aim of such educational technology is not to replace an instructor in DBT, but to supplement or inform users when an instructor is not present. This study makes no claims that virtual content is more effective than other media (for example, recorded videos of demonstrations). Instead this body of work is aiming to show the variety of ways in which mixed reality tools can aid education and training for DBT. By expanding the strategies to deliver observational content, we are providing effective learning environments for a broader audience.

Through the development process of the RPA, we can recommend these guidelines for delivering demonstration content in 3D virtual environments:

- Optimise interaction controls factoring the limitations of the target platform to achieve effective usability. The same application may need significant development to be effective on a different platform (desktop PC, tablet etc.)
- Autonomy of navigation and information flow is important to users and so may impact the perceived usability over long periods of usage. However, for short demonstrations, autonomy of information and navigation may not aid memory

recall, if the demonstration is paced appropriately and showcases the action sequence clearly.

- Information should be broken down into small steps to aid recall. The amount of information provided per step will depend on target user's prior knowledge and therefore this should be accounted for in the design process.

- Many mixed reality display types could facilitate virtual demonstrations. If the user has a variety of viewpoints to observe the required details, then a symbolic mental representation can be formed on both immersive, and non-immersive displays.

- Combining visual demonstrations with audible commentary is an effective tool for information delivery. Establish context through the environment details and generate demonstrator animations for all spoken details.

- In-application viewpoints should provide both egocentric and allocentric perspectives to aid the user's mental rotation of enacted movements and to support the use of varied vantage points to gather action details.

- Graphical treatment should clearly represent shape and form but does not need to be presented photo realistically to be effective. For demonstrations, prioritise animation fidelity over the shape and form of the performers and environment.

Future Directions

The participant feedback has highlighted additional features and refinements that would

improve the usability of the RPA navigation and information delivery:

- A repeat step function

- Breaking down of information into more steps

- Development of visual assets for context of the scene

- Adjustments to viewing angles so translational movement is not required

- Establishing visual actions or elements that represent all audible information

The next stage in the development of RPA will be to design and implement these features

for re-testing. Important to this will be to expand the participant pool in an ecologically valid

setting to see the extent to which users can physically perform observed movements. A

wider distribution and field testing with an instructor will be used to substantiate these

initial findings and inform future development of the application. There are also many more

mixed reality devices that may suit this type of demonstration. For example, augmented

reality (AR) devices where the virtual demonstration can be super imposed on a real-world

setting.

References

Akçayır, M. and Akçayır, G., 2017. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, pp.1-11. doi: 10.1016/j.edurev.2016.11.002

Ashford, D., Bennett, S.J. and Davids, K., 2006. Observational modeling effects for movement dynamics and movement outcome measures across differing task constraints: a meta-analysis. *Journal of motor behavior*, *38*(3), pp.185-205.

doi: 10.3200/JMBR.38.3.185-205

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice

Hall doi: 10.1177/105960117700200317

Boutin, A., Fries, U., Panzer, S., Shea, C.H. and Blandin, Y., 2010. Role of action observation and action in sequence learning and coding. *Acta psychologica*, *135*(2), pp.240-251.

doi: 10.1016/j.actpsy.2010.07.005

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). London, UK: Taylor & Francis. ISBN: 9780748404605

Brooks, B.M., Attree, E.A., Rose, F.D., Clifford, B.R. and Leadbetter, A.G., 1999. The specificity of memory enhancement during interaction with a virtual environment. *Memory*.

doi: 10.1080/741943713

Dalgarno, B. and Lee, M.J., 2010. What are the learning affordances of 3‑D virtual environments? *British Journal of Educational Technology*, *41*(1), pp.10-32. doi: [10.1111/j.1467-8535.2009.01038.x](10.1111/j.1467-8535.2009.01038.x)

Davis, S., Nesbitt, K. & Nalivaiko, E. (2014) 'A systematic review of cybersickness', In *Proceedings of the 10th 2014 Conference on Interactive Entertainment*, ACM, pp. 1–9. doi: 10.1145/2677758.2677780

Elliott, D., Grierson, L.E., Hayes, S.J. and Lyons, J., 2011. Action representations in perception, motor control and learning: implications for medical education. *Medical education*, *45*(2), pp.119-131. doi: [10.1111/j.1365-2923.2010.03851.x](10.1111/j.1365-2923.2010.03851.x)

Epstein, R.A., Patai, E.Z., Julian, J.B. and Spiers, H.J., 2017. The cognitive map in humans: spatial navigation and beyond. *Nature neuroscience*, *20*(11), p.1504. doi: [10.1038/nn.4656](10.1038/nn.4656)

Fery, Y.A. and Ponserre, S., 2001. Enhancing the control of force in putting by video game training. *Ergonomics*, *44*(12), pp.1025-1037. doi: [10.1080/00140130110084773](10.1080/00140130110084773)

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, Calif, Brooks/Cole Pub. Co.

Google Cardboard. (2018). *Google Cardboard – Google VR*. [online] Available at: https://vr.google.com/cardboard/ [Accessed 14 Aug. 2018].

Höffler, T.N. and Leutner, D., 2007. Instructional animation versus static pictures: A meta-analysis. *Learning and instruction*, *17*(6), pp.722-738. doi: https://doi.org/10.1016/j.learninstruc.2007.09.013

ISO 9241-11 (2018) Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts. Geneva: ISO. [online] Available at: https://www.iso.org/standard/63500.html [Accessed 20 Sep. 2018].

Jackson, P.L., Meltzoff, A.N. and Decety, J., 2006. Neural circuits involved in imitation and perspective-taking. *Neuroimage*, *31*(1), pp.429-439. doi: 10.1016/j.neuroimage.2005.11.026

Jang, S., Vitale, J.M., Jyung, R.W., Black, J.B., 2016. Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. Computers & Education. doi: 10.1016/j.compedu.2016.12.009

Krause, D., Kobow, S., 2013. Effects of model orientation on the visuomotor imitation of arm movements: The role of mental rotation. Human Movement Science 32, 314–327.

doi: 10.1016/j.humov.2012.10.001

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C. and Rizzolatti, G., 2005. Grasping the intentions of others with one's own mirror neuron system. *PLoS biology*, *3*(3), p.e79. doi: 10.1371/journal.pbio.0030079

Milgram, P. & Kishino, F. (1994) 'A taxonomy of mixed reality visual displays', *IEICE TRANSACTIONS on Information and Systems*, vol. 77, no. 12, pp. 1321–1329. [online] Available at: http://etclab.mie.utoronto.ca/people/paul_dir/IEICE94/ieice.html

Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), p.81.

doi: 10.1037/h0043158

Molenberghs, P., Hayward, L., Mattingley, J.B. and Cunnington, R., 2012. Activation patterns during action observation are modulated by context in mirror system

areas. *Neuroimage*, *59*(1), pp.608-615.

doi: 10.1016/j.neuroimage.2011.07.080

Norman, G., Dore, K. and Grierson, L., 2012. The minimal relationship between simulation

fidelity and transfer of learning. *Medical education*, *46*(7), pp.636-647.

doi: 10.1111/j.1365-2923.2012.04243.x

Pollock, E., Chandler, P. and Sweller, J., 2002. Assimilating complex information. *Learning*

and instruction, *12*(1), pp.61-86. doi: 10.1016/S0959-4752(01)00016-0

Rizzolatti, G. and Sinigaglia, C., 2010. The functional role of the parieto-frontal mirror circuit:

interpretations and misinterpretations. *Nature reviews neuroscience*, *11*(4), p.264.

doi: 10.1038/nrn2805

Rosen, M.A., Salas, E., Pavlas, D., Jensen, R., Fu, D. and Lampton, D., 2010. Demonstration-

Based training: A review of instructional features. *Human factors*, *52*(5), pp.596-609.

doi: 10.1177/0018720810381071

Sheffield, F. D. (1961) 'Theoretical considerations in the learning of complex sequential tasks

from demonstration and practice', *Student Response in Programmed Instruction*, pp. 13–

32. Doi: 10.17226/21290

Slater, M., 2003. A note on presence terminology. *Presence connect*, *3*(3), pp.1-5. Available

at:http://www0.cs.ucl.ac.uk/research/vr/Projects/Presencia/ConsortiumPublications/ucl_cs

_papers/presence-terminology.htm [Accessed 28 Sep. 2018]

Slater, M., 2014. *Grand challenges in virtual environments. Frontiers in Robotics and AI, 1,*

p.3. doi: 10.3389/frobt.2014.00003

Slater, M. and Sanchez-Vives, M.V., 2016. *Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3, p.74.* doi: 10.3389/frobt.2016.00074

Sweller, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), pp.257-285. doi: 10.1207/s15516709cog1202_4

Urgolites, Z.J. and Wood, J.N., 2013. Visual long-term memory stores high-fidelity representations of observed actions. *Psychological science*, *24*(4), pp.403-411. doi: 10.1177/0956797612457375

Van Merrienboer, J.J. and Sweller, J., 2005. Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review*, *17*(2), pp.147-177. doi: 10.1007/s10648-005-3951-0

Visbox.com. (2018). *Visbox, Inc. | CAVE Systems*. [online] Available at: http://www.visbox.com/products/cave/ [Accessed 20 Sep. 2018]

Vive.com. (2018). *VIVE™ | Discover Virtual Reality Beyond Imagination*. [online] Available at: https://www.vive.com/uk/ [Accessed 20 Sep. 2018]

Walker, M.P., Brakefield, T., Morgan, A., Hobson, J.A. and Stickgold, R., 2002. Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron*, *35*(1), pp.205-211. doi: 10.1016/S0896-6273(02)00746-8

Williams, A.M. and Hodges, N.J., 2005. Practice, instruction and skill acquisition in soccer: Challenging tradition. *Journal of sports sciences*, *23*(6), pp.637-650. doi: 10.1080/02640410400021328

Willingham, D.B., 1998. A neuropsychological theory of motor skill learning. *Psychological review*, *105*(3), p.558. doi: 10.1037/0033-295X.105.3.558

Wohlschläger, A., Gattis, M. and Bekkering, H., 2003. Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *358*(1431), pp.501-515. doi: 10.1098/rstb.2002.1257

Wulf, G., Schmidt, R. a., 1997. Variability of practice and implicit motor learning. Journal of Experimental Psychology: Learning, Memory, and Cognition 23, 987–1006. doi: 10.1037/0278-7393.23.4.987

Yang, Y., Leung, H., Yue, L. and Deng, L., 2013. Generating a two-phase lesson for guiding beginners to learn basic dance movements. *Computers & Education*, *61*, pp.1-20. doi: 10.1016/j.compedu.2012.09.006

Figures and Tables

Fig.1 Virtual Reality Examples: (Left) HTC Vive headset that is tracked in 3D space (Vive.com, 2018). (Right) A CAVE where the displayed image is projected onto a surrounding surface from the perspective of the user (Visbox.com, 2018).
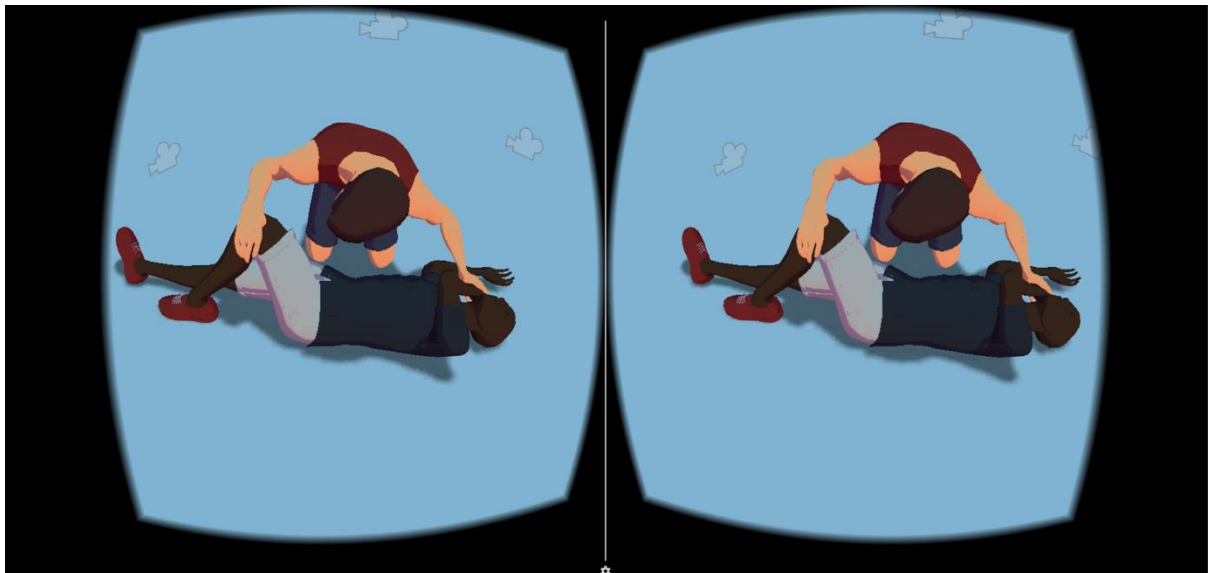

Fig.2. Shows how immersive displays render to each eye on the same display creating stereoscopic vision. Graphics of application are informative but optimised.

Fig 3. (Left) Shows how a mobile phone is placed as the display for Google Cardboard head mounted display. (Right) Shows the placement of the Google Cardboard display in use.
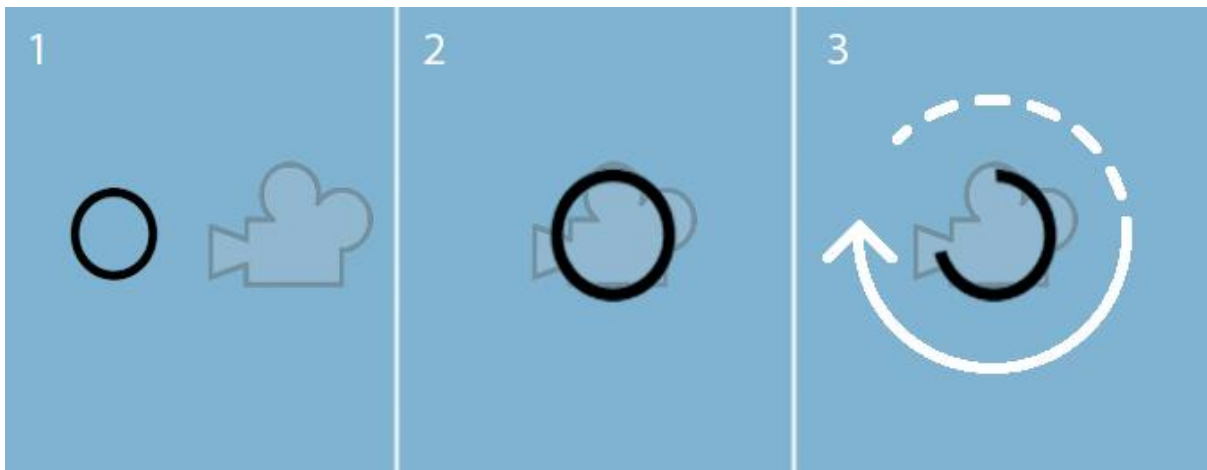

Fig.4. Image sequence showing time on target selection method of camera icon (panel 1). User reticle is positioned central to the viewport (Black circle, panel 1). User positions reticle over camera icon (panel 2). Reticle expands to provide feedback that icon is selectable. The reticle then redraws itself over 2 seconds (panel 3, white arrow shows direction of redraw). When the circle is fully redrawn, the item is selected.

Fig.5. Users can teleport to the camera locations (Marked "C") at any point during the demonstration. Demonstration will not progress to the next movement until users select the "Next Step" icon symbolised by a white circle (Marked "S").
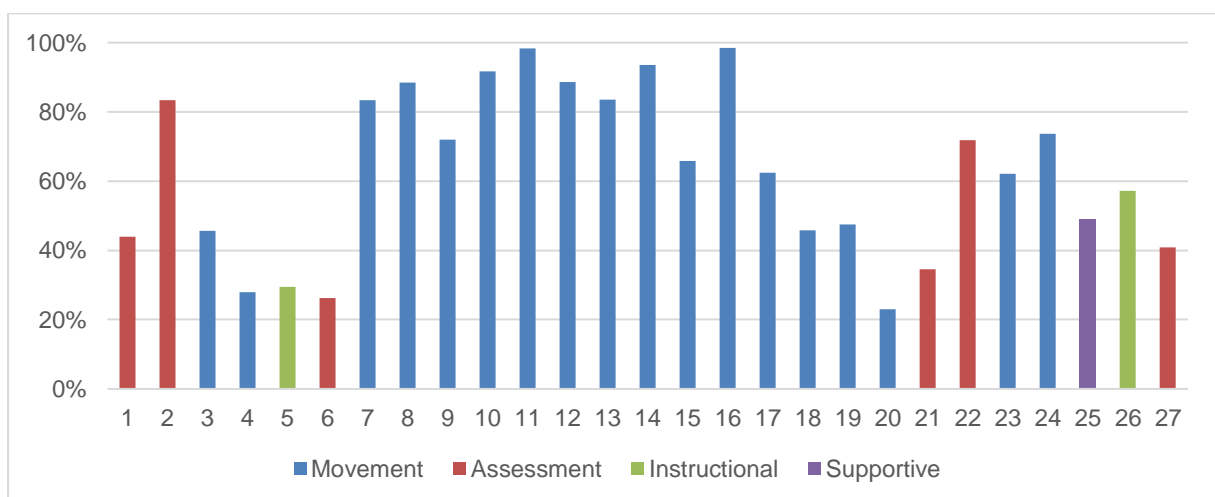

Fig.6. Average recall (across all groups) of details presented in the Recovery Position Application. These details are colour coded to the information type
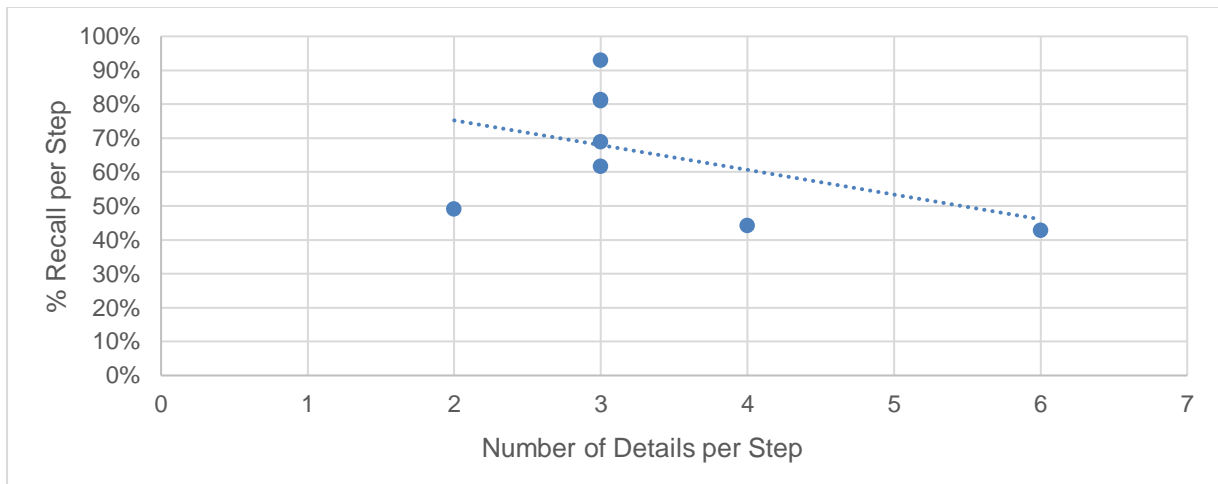
Fig.7. As the amount of details in each step increased, the % of information recalled
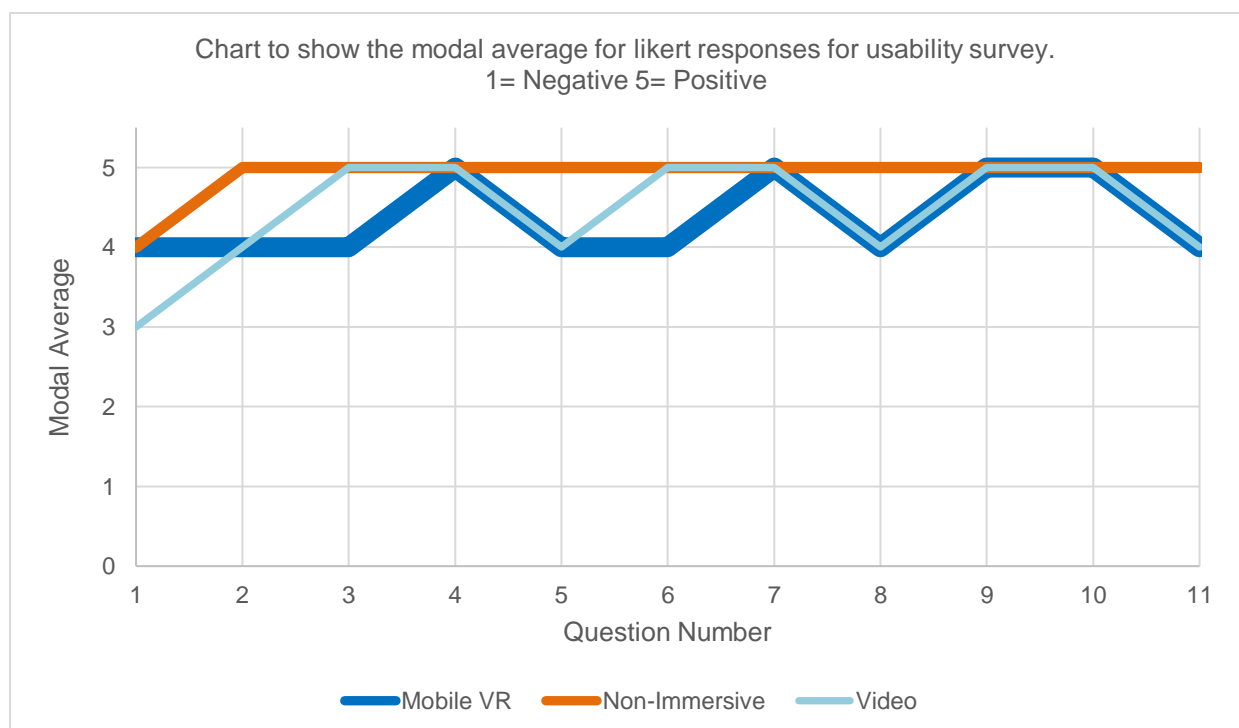
decreased.

Fig.8. Results of Likert questionnaire detailing participants' perception of the Recovery

Position application for all three groups. This indicates a general positive perception to the

usability of the application independent of the platform.

Information Delivered Through the Application

M = Movement, A = Assessment, I = Instructional, S = Supportive

Step No.	Detail No.	Detail Description	Info Type
01	1	Check the area poses no risk to yourself	A
	2	Check that the casualty is breathing	A
	3	Gently tilting the head back	M
	4	Listen and feel for breath on your cheek	M
	5	Look for movement in the chest	I
	6	Only proceed if they are breathing normally	A
02	7	Select arm closest to you	M
	8	Place at right angle to casualty's body	M
	9	Palm facing up	M
03	10	Select hand furthest from you	M
	11	Bring across casualty's body	M
	12	Place back of casualty's hand against patient's cheek	M
04	13	Grab knee furthest from you	M
	14	Raise it up	M
	15	Until foot is flat on the floor	M
05	16	Gently roll casualty towards you	M
	17	By pulling on the knee	M
	18	Support the casualty's head with your hand during this manoeuvre	M
06	19	Tilt the head	M
	20	By lifting the chin	M
	21	Ensure airway is open	A
	22	Check for normal breathing	A
07	23	Select top leg	M
	24	Bring out at right angle	M
	25	To support the casualty	S
08	26	Call an ambulance	I
	27	Monitor the casualty	A

Table 1. Table to show information delivered through the Recovery Position Application. The "Step No" of Table 1 categorises a group of details delivered in one section.

Memory Recall % for Information Types

Information Type	Mobile VR	Non-Immersive	Video	Average across conditions (Mean)
Movement Recall %	73.61	63.61	71.43	69.55
Assessment Recall %	47.50	43.33	59.52	50.12
Instructional Recall %	45.00	35.00	50.00	43.33
Supportive Recall %	30.00	50.00	67.67	48.89
Overall Recall % (Mean)	62.14	56.48	67.02	61.88

Table 2. Comparison between groups of recalled information categorised by information

type.