

RESEARCH

Open Access



Examining the validity and fidelity of a virtual reality simulator for basic life support training

Tom Arthur^{1,2*}, Tilly Loveland-Perkins², Charlotte Williams³, David Harris^{1,2}, Mark Wilson¹, Toby de Burgh², Jagtar Dhanda⁴ and Sam Vine^{1,2*}

Abstract

Background Virtual reality (VR) offers an immersive and practical method for training medical skills, especially in emergency healthcare settings. However, it is unclear whether learning in VR will translate into real-world performance benefits. To explore these potential transfer effects, we examined the validity and fidelity of a bespoke VR environment for Basic Life Support (BLS) training, a generic skill in medical training programmes.

Methods Twenty-two medical trainees performed standardised BLS procedures within two simulation conditions: one in VR, using a Pico Neo 3 standalone system; the other in a real-world synthetic environment, which included a physical mannequin and resuscitation equipment. Patterns of task behaviour, workload, sense of presence, and visual attention were derived from user self-report questionnaires, video recordings, and eye-tracking data.

Results Data showed that the VR training environment was sufficiently high in face validity to immerse the participants, and that trainees were displaying realistic task behaviours and procedural actions. However, the fidelity of user interactions and movements in VR proved atypical, which seemed to disrupt participants' attentional and motor responses.

Conclusions Results suggest that VR may have limitations for improving physical skills in the context of BLS training, yet be potentially valuable for developing task procedures and/or perceptual abilities.

Keywords Virtual reality, Simulation, Medical education, Healthcare

*Correspondence:

Tom Arthur

T.T.Arthur@exeter.ac.uk

Sam Vine

S.J.Vine@exeter.ac.uk

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Introduction

Simulation-based learning, using either virtual environments or synthetic physical equipment, is a key element of medical training. Indeed, the healthcare sector has often led the way in the development of evidence-based synthetic environments for enhancing occupational skills [1]. Here, learning typically occurs via physical simulation methods, in which trainees practice on human actors and real-life models or props [2]. For example, Basic Life Support (BLS) skills are routinely taught using a training mannequin like the 'Resusci-Anne', which can be physically interacted with and incorporated within controlled clinical scenarios. Such an approach offers a highly usable and safe method of practicing fundamental emergency procedures (e.g., the sequencing and/or performance of chest compressions, rescue breaths etc.) and has proven beneficial for both healthcare professionals and the general population [3–5]. These standardised procedures are, however, usually performed within organised teaching settings and are delivered by specialist training providers, making them inefficient and costly. Moreover, despite being relatively easy to implement, this type of training seldom recreates the actively changing sensory conditions and pressures that characterise 'real-world' BLS operations [6]. Crucially, these contextual differences can prevent trainees from experiencing authentic psychological and emotional responses during the learning process, which limits the overall effectiveness of a training programme [7]. As such, there is a demand for new immersive simulators that are both efficient and naturalistic for training purposes [8].

Virtual Reality (VR) technology provides new opportunities for simulating immersive sensory environments within medical education [9]. Commercial standalone headsets are becoming commonplace in healthcare, and their ability to create standardised clinical scenarios without the need for costly specialist staff or facilities makes them an increasingly appealing option for BLS training [10]. Studies have indeed shown that VR simulations can create highly immersive and well-accepted BLS learning environments that build skills, confidence and task knowledge in trainees ([11–14]; see [15] for recent review). Moreover, VR interventions have been found to enhance skills in a range of clinical domains, such as laparoscopic surgery [16], ophthalmology [17], neurological assessments [18], and mass casualty triage decision-making [19]. However, despite this promising initial support, caution must be placed when incorporating VR within occupational training programmes. Indeed, the effectiveness of training will not only depend on whether a simulation feels immersive, but also on whether it possesses the critical attributes for generating improvements in real-world performance [7]. Given that VR-related

differences in user behaviour and movement control could conceivably exist during BLS task procedures (e.g., during chest compressions: [20]), it is crucial that this potential transfer of learning is comprehensively evaluated before the technology is implemented within the field.

Harris et al. [7] have recently outlined two key predictors that can determine transfer of learning from VR. The first is *validity*, which refers to whether a simulated environment provides accurate representations of its real-world equivalent. To assess this component, researchers typically focus on a user's subjective view of how realistic a simulation is (i.e., face validity) and whether task-specific functionalities correspond between virtual and real-world conditions (i.e., construct validity: [21, 22]). The second predictor of learning transfer is *fidelity*, which concerns whether a simulation elicits realistic psycho-behavioural responses from its users [7, 23, 24]. This is assessed from a user's physical movements, affective state, and/or cognition. Indeed, though dependent on the specific objectives of training, the generalised nature of most clinical skills demand that simulations provide a suitable degree of realism at both a physical and psychological level [25]. From a BLS perspective, the likelihood of a VR simulator facilitating tangible skill improvements will therefore depend on whether it is representative of real-world tasks *and* capable of generating 'lifelike' user responses.

Based on the framework outlined by Harris et al. [7], the present study assessed the validity and fidelity of a VR simulator which has been designed to develop BLS competencies. To do this, we compared the VR simulation with a standardised real-world equivalent, which was based on current medical training practice in the UK. These two contrasting training conditions were closely matched for visuospatial features and presented within the same clinical context. Hence, we could evaluate the degree to which VR can provide an accurate representation of real-world sensory environments (e.g., by using self-report measures of user experience), and whether it elicits authentic psycho-behavioural responses (by measuring physical actions, cognitive workloads, and in-situ gaze responses). Crucially, these preliminary assessments were conducted *before* the simulation is implemented within training programmes (as recommended in [7]). This approach enabled us to not only gauge the potential efficacy of VR for enhancing clinical skills; it could also help guide additional software developments ahead of future training applications.

Given the positive learning outcomes reported in previous research (e.g., [11–14]), VR-based training

was hypothesised to create immersive, high-fidelity learning conditions. This would be reflected in authentic user experiences (i.e., high face validity) and positive correlations between clinical expertise and VR-based performance outcomes (i.e., high construct validity). From a fidelity perspective, patterns of cognition (e.g., perceived workloads) and behaviour (e.g., gaze and motor responses) were expected to be similar in VR to those shown in real-world operations. Such effects would indicate a strong potential for transfer of learning within the context of BLS training [7].

Methods

Participants

Twenty-two eligible medical trainees took part in the study (9 males, 13 females, age range: 20–38 years). These individuals had all received prior BLS training and presented varying levels of clinical proficiency and previous VR experience (see Table 1). Inclusion criteria stated that participants were undertaking a medical degree or clinical training programme within the UK National Health Service. Participants were excluded if they reported negative responses to VR, such as cybersickness or distress. The recruited sample size was sufficiently powered to

detect moderate-to-strong statistical effects in the data (i.e., between-condition effects equivalent to $d \geq 0.54$, at $p = 0.05$, $1 - \beta = 0.80$). All participants provided written informed consent, in accordance with British Psychological Society guidelines. The study was approved by the School of Sport and Health Sciences Ethics Committee, University of Exeter, and the experimental procedures adhered to this approved protocol at all times.

Simulation conditions

The simulated VR environment was presented using the Pico Neo 3 Pro Eye headset (Pico Interactive, San Francisco, CA): a lightweight, standalone head mounted display system with inbuilt eye tracking capabilities and a 98° field of view (refresh rate: 90 Hz). The device enabled participants to perform a BLS task within a virtual healthcare setting, while also recording pupil positions at 72 Hz. The VR environment was built using Unity game development software (version 2020.3.1; Unity technologies, CA) and was designed to simulate an empty hospital waiting room that users could freely move around in (see Fig. 1). Situated within this virtual room was a static mannequin, which replicated the half-body physical models that are used in real-world BLS training programmes. Additionally, the VR environment contained resuscitation equipment (a bag valve mask and three different-sized guedel airway devices), an emergency telephone, and a visible safety hazard (a wet floor sign and water puddle).

Participants could interact with the simulation objects by moving their virtual hands (the Pico controllers) to the object's 3-D spatial position and pressing a 'grip' button on the side of the controller. For instance, they would perform chest compressions by moving both controllers to the middle of the mannequin's torso, while holding down the grip buttons for the duration of each movement (see Supplementary Table 1 for a full list of the simulation task functionalities). An illustrative video of this

Table 1 Summary of descriptive measures from the study population

Mean Age in Years	26.48 (SD = 3.96)
Gender	9 male, 13 female
Current Healthcare Role	5 medical students, 5 foundation doctors, 10 SHOs, 2 registrars
Years of Completed Medical Training	6.83 (SD = 3.07)
Previous Virtual Reality Experience	7 with no prior experience of VR 7 with limited VR experience (< 2 h) 5 with some VR experience (2–20 h) 3 frequent standalone VR users

SD standard deviation, SHO senior house officer (i.e., doctors in post-foundation training or locally employed in non-training equivalent roles), VR virtual reality

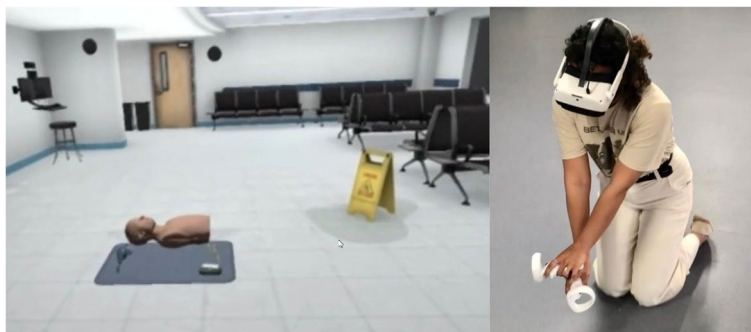


Fig. 1 The basic life support task for virtual simulation conditions. The figure on the left shows the user's view of the virtual scenario, and the illustration on the right shows the Pico Neo 3 Pro Eye headset and controllers that were used in the study

virtual environment and its functionalities can be found at <https://osf.io/eq4pc/>.

The virtual room was 147m² in total area, which was smaller than the surrounding laboratory workspace. Participants were able to move around the VR environment completely freely (e.g., by walking up to the mannequin, crouching on the floor etc.) and were not impeded by any obstacles or boundaries. Although the sizes of all objects in the VR environment were consistent with those used in real-world clinical settings, participants would not experience the physical sensations of interacting with these items (e.g., there would be no weight ascribed to the equipment or no resistance from the mannequin torso). Instead, user interactions were accompanied by vibrations of the hand controllers and representative auditory cues.

To ensure that participants were sufficiently accustomed to the artificial performance conditions, they completed a series of familiarisation tasks within a second VR environment. For these activities, participants were situated within a virtual dressing room area, where they would interact with two objects using the exact same methods as in the experimental study condition. Specifically, they were required to pick up a drinks can and put a baseball cap on their head using the VR hand controllers. These game items were presented on a long workbench in the middle of the room, at a distance that would require users to move around the virtual space.

The real-world conditions were performed in the exact same physical laboratory space as the VR and contained the same room layout and task objects (except for Guedel airway devices, which were not available in the real-world as the mannequin did not allow for insertion of airway adjuncts). The relative locations of the half-body mannequin, resuscitation equipment, and water hazard replicated those in the virtual simulation, to ensure that these key visuo-spatial features were identically matched between conditions. In the real-world task, participants wore Pupil Labs mobile eye tracking glasses (Pupil Labs, Sanderstrasse, Berlin, Germany), which recorded scene camera and pupil positional data at 90 Hz to provide an indication of dynamic gaze locations (spatial accuracy: 0.60°). Calibration of this system was performed before the task was initiated, using the manufacturers built-in screen marker routine.

Study design and procedures

Upon arriving to the laboratory, participants provided written informed consent and demographic information (as detailed in the *Measures* section). Thereafter, they would perform each of the study's two experimental conditions in a pseudo-randomised order. For the *VR condition*, participants were firstly fitted with the Pico headset

and introduced to the familiarisation environment. They were initially given up to one minute of exploration time in this simulation, in which they could freely move around the virtual space and make any necessary adjustments to the headset positioning (e.g., for comfort or enhancement of visual focus). During this time, participants were told that they would also be able to move around the BLS training environment in the same naturalistic and unconstrained way. Once comfortable with these task features, participants were then required to interact with the two familiarisation game objects (the baseball cap and drinks can), in any order of preference. These steps ensured that they were accustomed to the VR controls and functionalities ahead of the experimental BLS tasks. The familiarisation procedures were terminated once the participants had successfully interacted with both game objects and had verbally confirmed that they were ready to proceed with the main experimental tasks.

Before commencing the BLS task, participants received a standardised briefing from the researchers. These instructions conveyed situational information about the simulated environment, such as the cause of the emergency and the objectives of their intervention (see full scripts at <https://osf.io/eq4pc/>). Once participants had confirmed that they understood these instructions and were ready to proceed, the researchers initiated the task. Hereafter, participants were able to freely move around the virtual room and interact with the simulated patient (i.e., the half-body mannequin) and any resuscitation equipment. The task was deemed complete once three rounds of chest compressions and rescue breaths had been successfully delivered. At this point, and once the recording of all data outcomes had been saved, participants would take off the VR headset and then complete a series of self-report questionnaires.

For the *real-world* condition, participants were firstly fitted with the eye-tracking glasses and completed the standardised calibration procedures. They then received an identical briefing to the VR conditions and were shown to their initial position. Participants started 3.7 m away from the mannequin (as in the VR task), while facing the opposite direction from all task objects until the trial had commenced (to prevent goal-relevant visual cues from being retrieved before the onset of data recording). Once the task had been started, participants were instructed to turn on the spot before completing their subsequent BLS procedures. From this point, the real-world BLS task was exactly the same as the VR equivalent, both in terms of the background clinical scenario and the required behaviours. Once again, the trial was concluded upon the successful completion of three rounds of chest compressions and rescue breaths.

Crucially, neither experimental condition imposed any constraints or guidance on which specific BLS equipment or procedures should be operated by the participants. This was important, given the varying levels of clinical training and experience exhibited by the study sample (Table 1). Indeed, while some participants may have been less qualified or willing to employ certain procedures than others (e.g., rescue breath procedures using the bag valve mask), our repeated-measures analyses was only interested whether these key decision-related behaviours were *similar* between the VR and real-world simulations. As such, participants were simply informed that they should perform the BLS tasks in a manner that was consistent with their previous training. Upon completing both conditions, they were then debriefed by the researchers. Laboratory visits generally lasted under 45 min for each participant, and breaks were offered between each condition. All methods were performed in accordance with relevant guidelines and regulations.

Measures

Self-report measures

To examine face validity, we measured users' subjective sense of presence (i.e., the degree to which they felt as though they actually existed inside the VR environment) using a version of the Presence Questionnaire (adapted from [26]). This commonly used tool would illustrate whether the simulation was sufficiently accurate and realistic to create immersive user experiences [27]. Specifically, the questionnaire requires participants to respond to ten itemised statements on a 7-point Likert scale. Sub-item scores are then combined into an overall total, with higher scores signalling greater levels of presence. Values that exceed the mid-point of each scale would indicate that the participants were sufficiently immersed in the virtual environment and that the VR simulator was relatively high in face validity [28, 29].

To assess aspects of fidelity, we measured the psychophysical demands associated with each BLS training protocol using the Simulation Task Load Index (SIM-TLX; [30]). Participants completed this previously validated questionnaire after both simulation conditions, by self-rating levels of workload in nine separate items: mental demands; physical demands; temporal demands; frustration; task complexity; situational stress; distractions; perceptual strain; and task control. Each dimension was scored from 'very low' to 'very high' on a bipolar 21-point rating scale, with higher total scores signalling greater perceived workloads. The sum of the nine sub-item ratings were computed to provide a total SIM-TLX score for each participant.

Behavioural measures

All behavioural data were retrieved and processed offline, following inspection of task video recordings. These video recordings were obtained from a first-person perspective to facilitate the extraction of several performance metrics. In real-world conditions, the recordings were made by the Pupil Labs eye tracker's scene camera, which was positioned on the top of the glasses frame. For VR conditions, this footage was obtained from the simulator's customised remote viewing function, which displays user's point of view on a connected laptop. Using this footage, we were able to log each procedure that was undertaken by users, as well as the timing and frequency of key interactions. Specifically, we recorded the number of chest compressions and rescue breaths that were performed in each round and observed whether participants checked for consciousness, airway obstruction, breathing, and circulation in their simulation (in accordance with Resuscitation Council UK guidelines). For these binary event-related outcomes, we assigned a score of 1 for actions that were undertaken by participants and a score of 0 in instances when the actions were not performed.

Moreover, we calculated the time taken to perform the BLS task in each of the study's simulation conditions. Video recordings were manually inspected in a frame-by-frame fashion to calculate the elapsed time between the onset of the task and the successful completion of three rounds of chest compressions and rescue breaths. Taken together, these measures would indicate the degree to which task behaviours in VR correspond with real-world performance actions and expertise [7, 21, 22].

Eye-tracking measures

To further assess aspects of fidelity, we compared users' visuomotor responses between the two study conditions. Indeed, the continuous regulation of gaze during movement-based tasks, coupled with the sampling of goal-relevant sensory information, can provide objective indicators of clinical expertise (e.g., [31, 32]), decision-making biases [33], emotional regulation (e.g., [34, 35]) and perceived workloads [36, 37]. Hence, potentially meaningful differences in visuomotor behaviour could be extracted from user's dynamic gaze responses. Specifically, both eye tracking systems that were used in the current study produced combined gaze vector positions in cartesian (x, y, z) coordinates. These raw datafiles were first inspected for signal quality and then analysed using customised MATLAB scripts (available at <https://osf.io/eq4pc/>). To enable comparisons between conditions, positional data were converted into angles on an equivalent 'gaze-in-head' spherical coordinate system (i.e., phi, theta, and radius values, relative to head orientation).

Thereafter, the angular coordinates were resampled to a consistent 36 Hz and passed through a zero-phase Butterworth filter (at 15 Hz for positional data and 50 Hz for velocity data, as in [38]). From here, a number of key gaze metrics were calculated, as described below.

Saccade frequency Rapid shifting of gaze to a new visual location (i.e., saccades) were identified from portions of gaze data that exceeded five times the median acceleration [39]. Gaze velocity had to be at least 30°/s during this period of time, and over 15% of the trial-specific maximum velocity [40, 41]. Any data that were preceded or followed by missing values were disregarded, to avoid erroneous detections. The number of detected saccades were then divided by the total task duration to provide a relative frequency value (i.e., saccades per second). Higher frequencies signalled that participants were shifting their gaze more readily around their surrounding visual workspace.

Saccade type To further understand participants visual search behaviours, the change in angle between successive saccades was calculated to classify persistent and antipersistent strategies [42]. Persistent saccades were detected from eye movements that shifted gaze continuously in a direction that was within 90° of space. Conversely, antipersistent saccades were those that changed direction by >90°. The proportion of each type of saccade was expressed as a percentage, before being compared between conditions. A higher proportion of antipersistent saccades within a given condition would illustrate a large amount of inefficient ‘back and forth’ gaze shifts, whereas a high percentage of persistent saccades would reflect smoother, more continuous visual scanning patterns across the simulated workspace.

Average fixation durations Fixations were defined from clusters of gaze data that fell within 1° of visual angle for a minimum of 100 ms [43]. The duration of each fixation event was made using a well-established spatial dispersion algorithm [44], before being averaged for each participant in each condition. Longer average durations within a task condition would indicate prolonged sampling of visual cues.

Entropy To assess how structured or efficient participants’ visual search behaviours were, we assessed Gaze Transition Entropy [45]. This measure indexes levels of variability or randomness in the continuous eye tracking data, by calculating the probability of a given data-point (i.e., current fixation location) being conditional upon previous recorded values (i.e., preceding fixation locations). To categorise our gaze-in-head positional

data, the egocentric visual scene was split into 15 content-independent areas of interest (AOIs), based on a uniform 5 × 3 grid. The AOI grid followed dimensions that are consistent with previously reported studies (e.g., [46]). Specifically, for both phi and theta coordinates, central segments represented fixations that were ≤ 12.5° from the midpoint of the visual scene. On the phi axis, fixation locations that were <25° to either side of this central AOI represented the next layer of AOIs, while those that deviated from the midpoint by >37.5° were assigned to outer (peripheral) segments. For theta coordinates, the outer segments represented gaze locations >12.5° from the scene midpoint (i.e., values that were above or below the central segment). After assigning each fixation to an AOI, entropy was calculated using the following equation:

$$Entropy = \sum_{i=1}^n p(i) \left[\sum_{j=1}^n p \left(\begin{matrix} j \\ - \\ i \end{matrix} \right) \log_2 p \left(\begin{matrix} j \\ - \\ i \end{matrix} \right) \right], i \neq j$$

Here, the sum of the logarithm of all conditional probabilities (which signifies the likelihood of fixating each AOI) is estimated for a given state space in ‘bits’, with *i* representing preceding gaze locations and *j* representing the next location in the sequence. In sum, when gaze is shifted predictably between strategic and regular locations in space, entropy will be relatively low; but when visual search behaviours follow erratic and reflexive patterns over time, then entropy will be relatively high.

Statistical analysis

Data outcomes were initially screened for missing and/or extreme values (*p* < 0.001), and for any extreme deviations from normality, linearity, multicollinearity, or homoscedasticity. Univariate outliers were Winsorised to 1% larger or smaller than the next most extreme score. The cleaned data variables were then assessed for between-condition differences, using a series of paired *t*-test (for parametric data) or Wilcoxon-signed rank test (for non-parametric data) comparisons. Here, any discernible differences in gaze behaviour between virtual and real-world conditions would indicate that the VR simulator is not fully representative of real-world BLS environments and that it is eliciting atypical user responses. Conversely, a lack of between-condition differences would signal that the VR simulator is high in fidelity, and that it is more likely to facilitate transfer of learning within BLS training [7].

To examine levels of concurrent validity, Spearman's Rho analysis studied relationships between prior clinical expertise (number of years in formal medical training) and each of our continuous behavioural and eye tracking metrics. Here, significant positive correlations would indicate that the VR training simulation is sufficiently representative of 'real-world' medical proficiencies (as in [7, 21, 22]).

All statistical tests were performed using JASP (version 0.16.3) and are reported alongside a Bayes Factor computation (BF_{10}), which indicates the strength of evidence in favour of a null versus the alternative hypotheses (in accordance with [47]). Significance was accepted at $p < 0.05$ and averages are presented alongside a relevant standard deviation (SD) value. The study's full anonymised dataset is freely available at <https://osf.io/eq4pc/>.

Results

Preliminary data analyses

One participant displayed symptoms of cybersickness, meaning that they were excluded from the study. A further three participants were excluded from eye tracking analyses, due to missing data or poor tracking quality. This afforded a final sample of 21 for our self-report and behavioural data analyses, and a sample of 18 for our between-condition gaze data comparisons. Behavioural data relating to the number of chest compressions and rescue breaths in each simulation were disproportionately clustered around guideline values (i.e., 30 chest compressions and 2 rescue breaths per round/cycle of cardiopulmonary resuscitation; in line with recommendations made by the Resuscitation Council UK). Moreover, participants' time to task completion and average fixation durations were positively skewed. As such, these outcome

measures were analysed using non-parametric statistical procedures. No other deviations from normality, linearity, multicollinearity, or homoscedasticity were observed.

Self-report data

Scores from the presence questionnaire were relatively high following performances in the VR task conditions. Mean total values of 43.24 ± 6.92 exceeded the mid-point of the itemised scale (i.e., 40), suggesting that users felt like they really existed in the virtual environment. In fact, 15 participant totals (71.48%) were above this threshold, indicating that the high feelings of presence were widely prevalent in the study.

SIM-TLX scores significantly differed between virtual and real-world conditions ($t(20) = 9.01$, $p < 0.001$, $d = 1.97$; $BF_{10} = 7.43 \times 10^5$). As illustrated in Fig. 2, the VR task simulation was perceived to be considerably more demanding than the traditional mode of BLS training, an effect underpinned by elevated mental demands, frustration, complexity, distractibility, perceptual strain and difficulties with task control. Interestingly, though, Fig. 2 shows that the VR simulation was deemed to be substantially *less* physically demanding by users than its real-world equivalent.

Behavioural data

The average number of chest compressions (VR: 32.18 ± 4.11 ; real-world: 30.84 ± 1.93) and rescue breaths (VR: 1.94 ± 0.42 ; real-world: 2.11 ± 0.39) for the overall sample was not significantly different between conditions (p 's > 0.05 ; $BF_{10} < 1$), and the proportion of users who inspected the patient for consciousness (85.71%) and airway obstruction (95.24%) was identical in both environments. Moreover, despite being free to employ

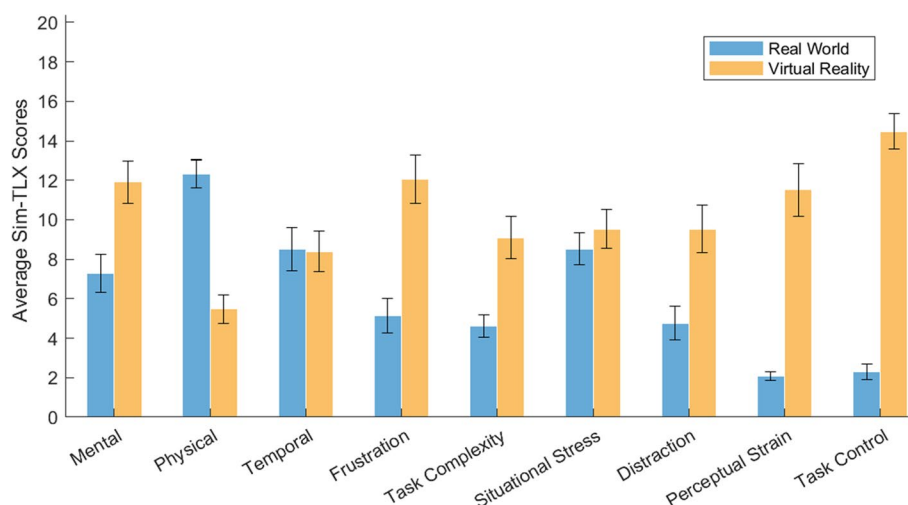


Fig. 2 Self-reported scores on the Simulation Task Load Index (SIM-TLX) during real-world and virtual reality conditions. Values represent averages for each sub-item on the questionnaire and error bars denote standard errors of the mean

either mouth-to-mouth or bag valve mask rescue breath techniques, the proportion of participants who utilised each method did not significantly differ between conditions ($\chi^2 = 0.171$, $p = 0.68$). Specifically, 88.24% (15/17) of participants that opted to use the bag valve mask in VR also opted to do so in the real-world. This suggests that users were generally undertaking similar procedures in the two distinctive training environments. However, trainees took significantly longer to perform the task when it was simulated in VR (Wilcoxon Signed-Rank: $Z = 4.02$, $p < 0.001$, $BF_{10} = 635.28$). Indeed, the time taken to complete three cycles of cardio-pulmonary resuscitation was 94.64% higher in the VR compared to the real-world training conditions (see Fig. 3).

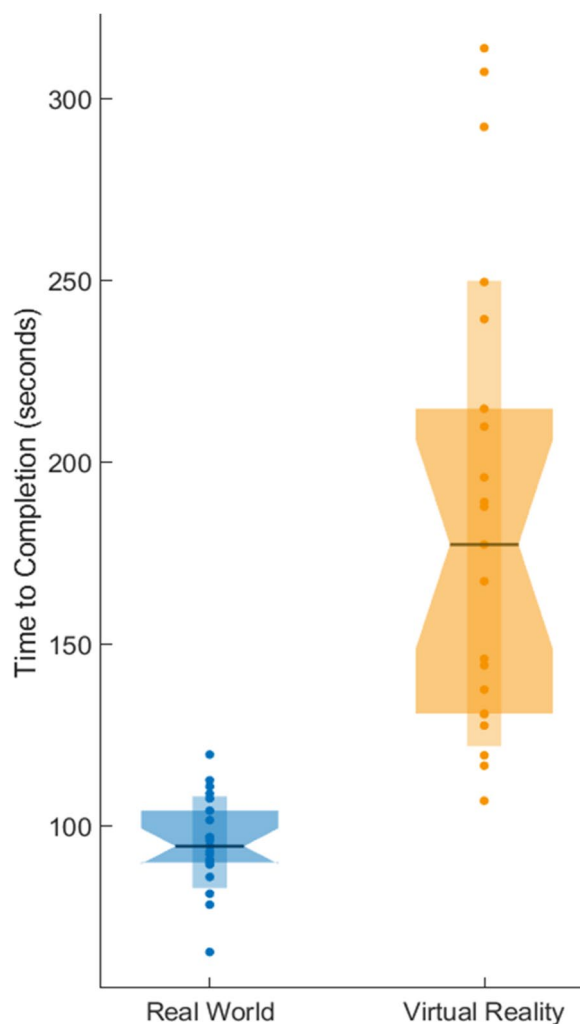


Fig. 3 Time to complete the task in real-world and virtual reality simulations. Horizontal lines represent the median for each study condition and the notched shaded areas denote upper and lower quartiles. Rectangular shaded areas depict the standard deviation. Filled circles represent individual cases

Eye-tracking data

There were some consistent patterns observed in the ‘real-world’ eye-tracking footage. When commencing the task, participants tended to initially scan across the room via a series of large saccades and successive fixations. This rapid sampling of visual cues enabled key task-relevant information to be retrieved from the scene, such as the existence of any safety risks (e.g., the wet floor hazard) and assistive support (e.g., an available helper or defibrillator). Participants often repeated these search behaviours when performing chest compressions, although their gaze was also sometimes directed to action-focused cues (e.g., towards the mannequin’s torso). Such ‘anchoring’ of gaze became more prominent during the provision of rescue breaths, when participants would tend to alternate their focus between the facial attachment of the bag valve mask and the middle of the mannequin’s torso. These strategic, goal-driven gaze responses are illustrated in Supplementary Videos at <https://osf.io/eq4pc/>.

Notably, participants displayed a reduced frequency of saccadic eye movements within the VR environment ($t(17) = 2.8$, $p = 0.01$, $d = 0.66$, $BF_{10} = 4.45$; Fig. 4A), which illustrates that gaze was being shifted less readily around the visual workspace. The type of saccades being used also proved to be atypical, as users exhibited a lower proportion of persistent saccades and a higher proportion of anti-persistent saccades under VR conditions ($t(17) = 5.60$, $p < 0.001$, $d = 1.32$, $BF_{10} = 781.46$; Fig. 4B). This suggests that the large, continuous visual scans that were prevalent in real-world BLS training environment were less prominent in the VR simulation, and that participants were instead relying on less efficient ‘back and forth’ gaze shifts within this setting.

Fixation behaviours also proved different between training conditions. For example, average durations were significantly shorter in the real-world (mean = 0.21 ± 0.05 s) compared to VR (mean = 0.25 ± 0.05 s; Wilcoxon Signed-Rank test: $Z = 2.81$, $p < 0.01$, $BF_{10} = 9.22$; Fig. 4C). This signals that participants were sampling virtual cues for longer than their real-world sensory equivalents, and that the control of visual attention may have been atypical during VR trials. Moreover, when analysing the structure and/or variability of participants’ fixation behaviours, results showed that gaze transition entropy values were significantly higher in VR compared to the real-world simulation environment ($t(17) = 6.00$, $p < 0.001$, $d = 1.42$, $BF_{10} = 1604.10$). This indicated that gaze shifts were less systematic and predictable within a VR setting.

Relationships with clinical expertise

There were no significant correlations detected between years of previous medical training and any of the

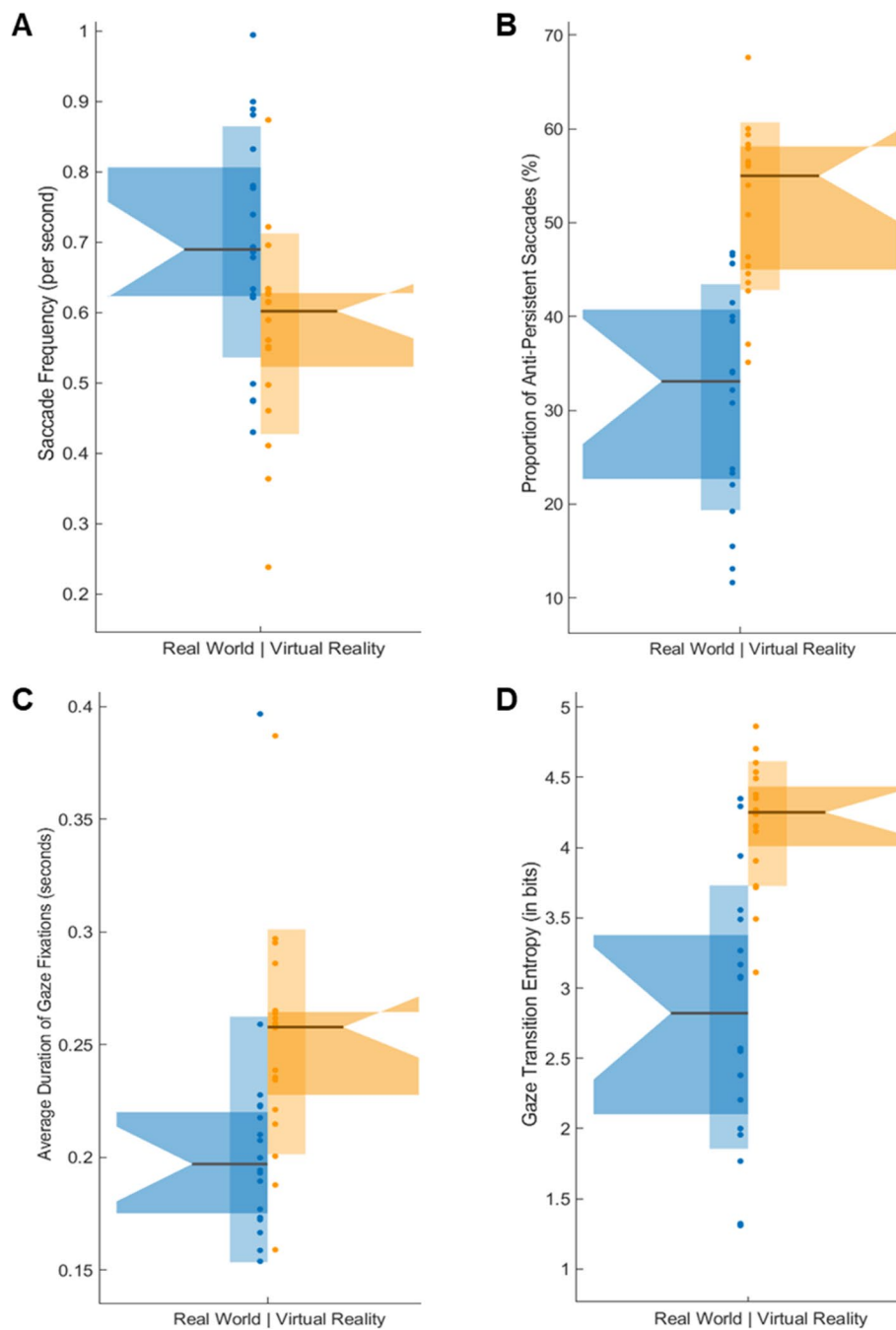


Fig. 4 Gaze patterns in the real-world and virtual reality simulations. Significant between-condition differences are illustrated for saccade frequency (A), the proportion of anti-persistent saccades (B), average fixation durations (C) and gaze transition entropy (D; all $p < .01$). Horizontal lines represent the median for each condition. Notched shaded areas denote upper and lower quartiles. Rectangular shaded areas depict the standard deviation. Filled circles represent individual cases

continuous workload, behavioural or gaze metrics in this study (p 's > 0.25 , $BF_{10} < 1.12$; see Fig. 5). This highlighted that prior clinical experience was unrelated to BLS task performance in either simulated training condition.

Exploratory analysis

The markedly different gaze responses in Fig. 4 could be explained by two hypotheses:

i) Users could be processing virtual cues in a fundamentally different way from those in the real-world. This

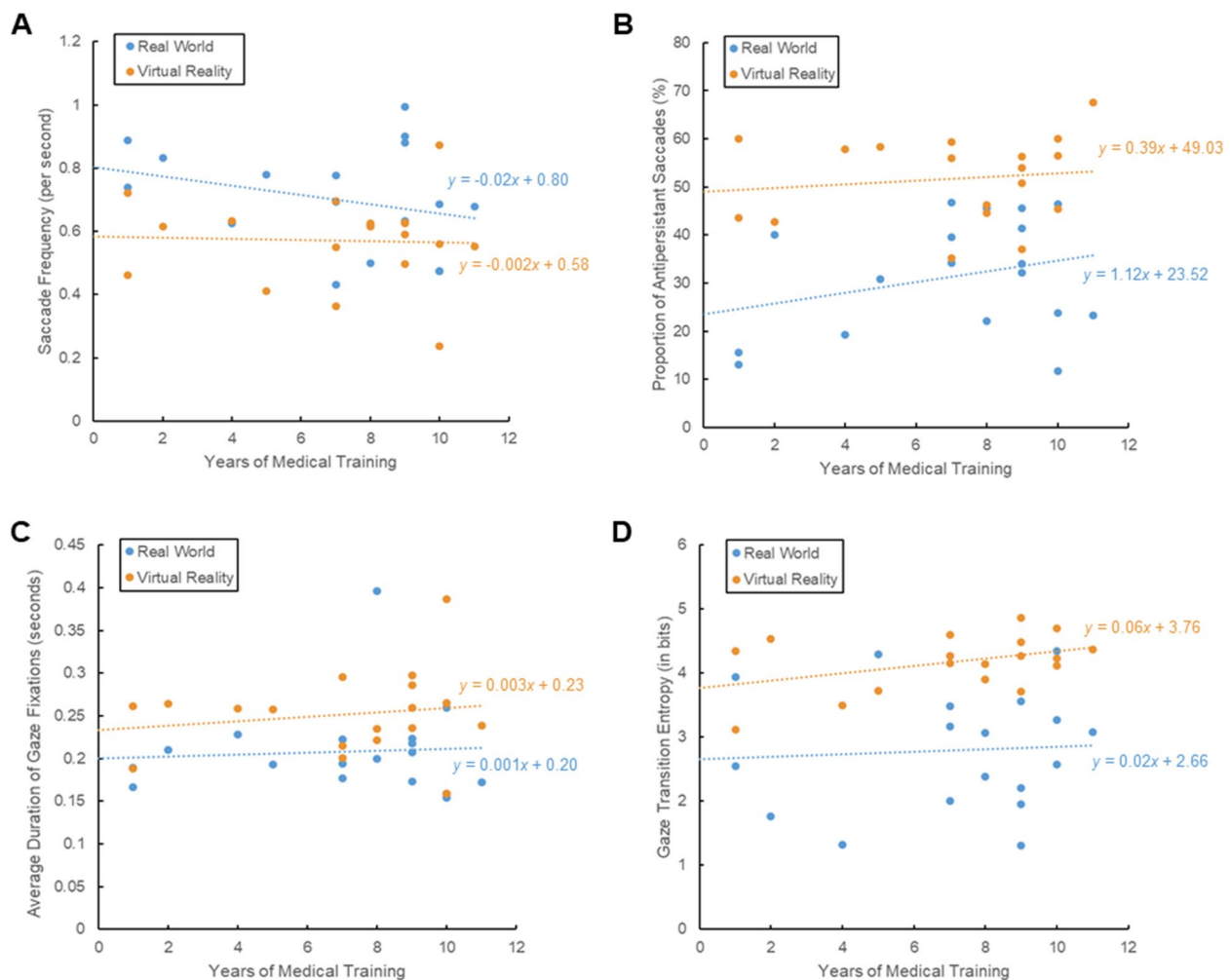


Fig. 5 Scatter plots highlighting relationships between years of previous medical training and gaze behaviours during the present study. Null statistical associations emerged for saccade frequency (**A**), the proportion of anti-persistent saccades (**B**), average fixation durations (**C**) and gaze transition entropy (**D**) during both real-world and virtual reality training environments (all $p > .05$). Dotted lines represent the line of best fit for each study condition. Filled circles represent individual case values

is consistent with observations that the integration of multisensory cues differs under conditions that are more uncertain or unrelated to prior experience (Kording et al., 2007).

ii) Conversely, altered gaze patterns could relate to differences in visuomotor control. Indeed, the VR task was deemed less physically demanding, but more complex and frustrating to perform (Fig. 2). Participants also took longer to successfully complete the simulated VR procedures (Fig. 3). As such, the usually efficient and automatic control of sensorimotor actions may have been disrupted in VR, leading to an atypical use of visual feedback cues.

Since these two hypotheses present divergent implications for clinical training, we analysed gaze behaviours during initial phases of the BLS task (when movement

demands were low and various perceptual assessments were instead being made). Specifically, we tested whether saccadic frequency and fixation durations varied between VR and real-world conditions *prior to the onset of any chest compressions*. As shown in Fig. 6, these outcomes did not significantly differ between training conditions (p 's > 0.19 , $BF_{10} < 0.53$). Therefore, it appears that simulation fidelity was relatively high in VR when the BLS task consisted of mostly perceptual components (e.g., when users initially assess the situation to determine an appropriate course of action), but low when the task involved dynamic motor actions and movements (e.g., during the provision of chest compressions and rescue breaths).

Crucially, this analysis shows that the prolonged fixation durations and reduced saccade frequencies that

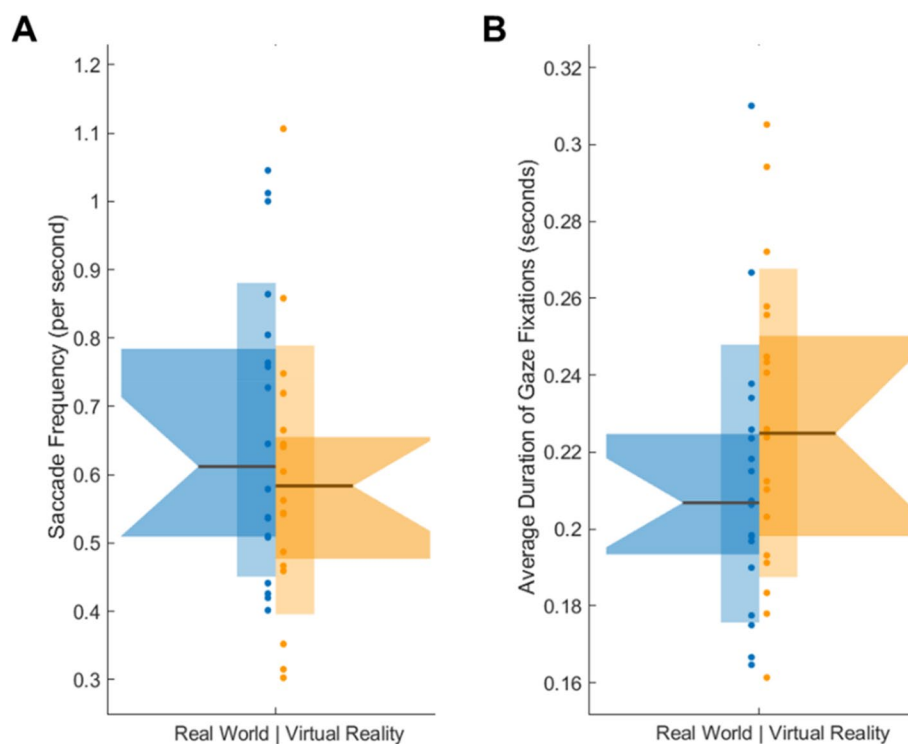


Fig. 6 User gaze patterns in the real-world and virtual reality training environments before the onset of chest compressions. Horizontal lines represent the median for each study condition and the notched shaded areas denote the upper and lower quartiles, whilst the rectangular shaded areas depict the standard deviation of the mean. Filled circles represent individual cases

were displayed in VR during the extended BLS task do not appear related to any generic abnormalities in the processing of virtual sensory cues. Instead, they likely reflect a more feedback-driven mode of visuomotor control, as consistent with Harris et al. (2019). While users in real-world training conditions were able to control their movements without needing to continuously monitor their actions (allowing them to frequently scan around the scene for alternative situational cues), they seemed to increasingly rely on incoming visual cues in VR. As a result, their gaze behaviours were more reflexive and action-focused in these training conditions.

Discussion

VR technologies could provide an appealing method for delivering BLS training (see [15, 48]). However, the degree to which these new and immersive forms of training foster practically meaningful learning effects (which transfer onto real-world performance) remains unclear. Consequently, we focused on two key predictors of skill transfer – simulation validity and fidelity – to investigate the potential utility of VR in the context of BLS training. Through integrating self-report user feedback with objective behavioural and eye-tracking data, our analysis presents some notable strengths and limitations of VR-based

methodologies in this field. Such features should be considered when designing future simulation training interventions.

Firstly, to evaluate simulation *validity*, we assessed whether the VR task provided accurate and immersive conditions for our user group of medical trainees. Participants generally reported feeling high levels of presence, with mean questionnaire scores exceeding those documented in other occupational domains (e.g., aviation: [28]). These data not only signal that participants felt like they really existed in the VR environment; they also support previous findings that VR can simulate immersive and realistic BLS learning conditions from the perspective of its users [11–13]. Such high levels of simulation validity are an important determinant of effective skill transfer [7] and may contribute to more adaptive behaviours and task motivation during learning [49]. Our results therefore reinforce the notion that VR could offer an engaging and immersive method of teaching BLS skills.

Our second criterion for evaluating validity was to examine whether VR accurately captured individual differences in task expertise. Results provided no support in this regard, with years of prior medical training proving unrelated to all of our study measures. The reasons for

these null effects could be twofold. Firstly, it is possible that the simulation did not provide sufficient construct validity. Indeed, if a training method does not accurately represent the functional parameters of real-world conditions, then it is unlikely to produce expert-related variations in behaviour [7]. However, null correlations were detected for both VR and real-world conditions in our data. So, one must also consider that there may not have been sufficient variability or sensitivity in our measures of task expertise to detect a relationship. The fact that all participants had received previous BLS training indicates that there may have been a 'ceiling effect' in the data. This is supported by our behavioural observations, which showed very few detectable 'errors' being made in either condition. To progress this research in the future, studies may therefore wish to examine how BLS task behaviours change in novice trainees over time, following repeated practice in VR.

When inspecting outcomes relating to simulation fidelity, our data show mixed results. From a behavioural perspective, we found that users reported higher perceived workloads and took significantly longer to perform the task in VR. This impeded delivery of cardiopulmonary resuscitation could be potentially detrimental in a clinical setting, since the rate of chest compressions and 'time off the chest' are considered key predictors of positive patient outcomes (e.g., see Resuscitation Council UK guidelines). Our eye-tracking data also implied that learners were sampling virtual sensory cues very differently from those in the real-world (Fig. 4). For instance, participants shifted their attention less frequently and predictably around the VR workspace (as indicated by lower saccadic frequencies and gaze entropy). Moreover, instead of employing the highly systematic visual scan behaviours that were displayed in the real-world simulation, gaze was increasingly shifted 'back and forth' in VR and held steady on cues for longer fixation durations. Given the clear difficulties that some users experienced when performing movement-based cardiopulmonary resuscitation actions in VR, as well as the limited haptic information that was made available in the simulation, we speculate that such a response is likely related to disruptions in visuomotor control. Indeed, atypical gaze responses were not present prior to the onset of chest compression actions in this task (Fig. 6) and atypical cardiopulmonary resuscitation movements have also been documented in previous VR studies (e.g., [20]). Furthermore, motor learning research has shown that learners can rely on suboptimal movement strategies and perceptual cues when interacting with virtual environments [50, 51]. Thus, when taken together, our results suggest that the VR simulation was lacking in aspects of physical and/or ergonomic fidelity, which is likely to have impacted

on the attentional and cognitive responses that were displayed by users.

Nevertheless, there were aspects of simulation fidelity that were more encouraging. For instance, users generally undertook the same clinical actions and decisions in the virtual simulation as they did in the real-world conditions, with Bayes factors for numbers of chest compressions and rescue breaths favouring the null model. Participants also showed realistic gaze responses during initial stages of the VR task (Fig. 6). Crucially, the initial phases of BLS consist of various situational assessments, whereby responders are required to actively check the state of both their patient and their surrounding scene. During these instances, participants employed wide-ranging systematic scanning procedures, which enabled the sampling of various visual cues from across the workspace. The fact that users performed these procedures comparably between VR and real-world conditions suggests that sufficient levels of psychological fidelity may have been achieved during parts of the VR simulation. Indeed, research has demonstrated that visual search abilities can be readily enhanced using VR-based learning methods [52]. Therefore, while the artificial task constraints and user mechanics in VR may have disrupted the regulation of movement-based procedures, some of the perceptual components of BLS appeared to remain intact and may thus be 'trainable' in the future.

Limitations and future research

A number of limitations must be considered prior to the implementation of VR training in the field. In particular, our approach of comparing VR with an equivalent real-world simulation (and not actual performance) must be acknowledged, as the degree to which users responded to our 'control' task in a truly realistic manner is ultimately unclear. Indeed, it is entirely possible that user responses in VR were more representative of actual BLS operations than the ones displayed under the simulated 'real-world' conditions. The use of the 'Resuci-Anne' method does, however, represent the current best practice (i.e., gold standard) of BLS training in the UK, so was a relevant comparison for our pre-implementation evaluation. Moreover, the highly consistent findings that emerged across users' self-report, behavioural and gaze data remain effective in highlighting areas of strength and limitation for future training tools.

Nevertheless, the present research did not include any direct measures of task performance. Given that all participants in this study were amply competent at undertaking the relatively simplistic BLS procedures, there would have been little value in attempting to scrutinise any minor, potentially trivial inter-individual

differences in motor proficiency (especially since movement atypicalities could yet exist in real-world simulation conditions). That said, future work could yet exploit the unique potential that VR methodologies afford in this domain (for further discussion, see [48]). For instance, researchers could evaluate whether VR software can automatically detect markers of successful and/or errorful behaviours in novice populations. Conversely, they could adapt the simulations to introduce more complex and/or stressful task conditions for expert user populations, through the use of challenging and individualised clinical scenarios (see [53] for specific examples). It is recommended that future studies are conducted with larger, more diverse sample populations, so that potentially significant individual differences and correlations can be explored more comprehensively.

Conclusions

Overall, this study suggests that VR-based simulation methods may be limited for improving visuomotor skills in the context of BLS training, but potentially valuable for developing transferable perceptual and/or procedural abilities. Results showed that our VR simulation was sufficiently accurate and immersive to make a group of experienced medical trainees feel ‘present’ and perform naturalistic procedural assessments. However, the fidelity of movement-based interactions proved limited, leading to higher self-reported workloads, longer times to task completion, and disrupted attentional responses. Although the fidelity of such interactions could be enhanced by new technological advancements in the field (e.g., improved hand tracking capabilities and haptic feedback), our results support further investigations into the use of different forms of simulation training for enhancing different aspects of BLS performance, and more general medical skills.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s44247-023-00016-1>.

Additional file 1: Supplementary Table 1. A list of procedures that were simulated in the Virtual Reality task.

Acknowledgements

The authors would like to thank all of the participants who took part in this study. We would also like to thank ExR for their design of the virtual reality simulation and Emteq Labs for their collaborative efforts and provision of technical support/equipment.

Authors' contributions

T.A., C.W., D.H., M.W., T.dB., J.D., and S.V. contributed in the design of this study. T.A. and T.L.P. performed data collection and analysis. D.H. provided analysis tools and assistance. T.A., C.W., D.H., M.W., J.D., and S.V. contributed in the writing of the paper. The author(s) read and approved the final manuscript.

Funding

This work was supported by a grant from Health Education England. Health Education England did not play any role in the design of the study, collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

Relevant data and code from this study are publicly available at: <https://osf.io/eq4pc/>

Declarations

Ethics approval and consent to participate

All methods were performed in accordance with the relevant guidelines and regulations. The study received ethical approval on 7th March 2022 from the School of Sport and Health Sciences Ethics Committee (University of Exeter, UK). All participants were sent an information sheet prior to taking part which detailed the research aims, methods, availability of results, confidentiality of data, and their rights as participants. On arrival to the laboratory, they were given the opportunity to ask any further questions before providing written informed consent (in accordance with British Psychological Society guidelines).

Consent for publication

Participants provided written informed consent for their anonymised data to appear in academic publications and for it to be uploaded onto an associated online repository for further analyses. All study data were de-identified by the research team accordingly, to adhere to these approved research protocols.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Public Health and Sports Sciences, Faculty of Health and Life Sciences, University of Exeter, Heavitree Road, Exeter EX1 2LU, Devon, UK.

²Cineon Training Ltd., Exeter Science Park, 6 Babbage Way, Clyst Honiton, Exeter EX5 2FN, UK. ³Bristol Medical School, Faculty of Health Sciences, University of Bristol, Bristol BS8 1TH, UK. ⁴Virtual Reality in Medicine and Surgery (VRiMS.net), Brighton and Sussex Medical School, BN19PX Brighton, UK.

Received: 14 December 2022 Accepted: 20 April 2023

Published online: 11 May 2023

References

1. Drews FA, Bakdash JZ. Simulation training in health care. *Rev Hum Factors Ergon.* 2013;8(1):191–234.
2. Alanazi AA, Nicholson N, Thomas S. The use of simulation training to improve knowledge, skills, and confidence among healthcare students: a systematic review. *Internet J Allied Health Sci Pract.* 2017;15(3):2.
3. Banasik Z, Sledziński Z, Arciszewska D, Wawel M, Kucharska K, Lewiński A. The usefulness of Resusci-Anne manikin in teaching modern methods of resuscitation. *Anaesth Resusc Intensive Ther.* 1976;4(2):131–7.
4. García-Suárez M, Méndez-Martínez C, Martínez-Isasi S, Gómez-Salgado J, Fernández-García D. Basic life support training methods for health science students: a systematic review. *Int J Environ Res Public Health.* 2019;16(5):768.
5. Wanner GK, Osborne A, Greene CH. Brief compression-only cardiopulmonary resuscitation training video and simulation with homemade mannequin improves CPR skills. *BMC Emerg Med.* 2016;16(1):1–6.
6. Wisborg T, Brattebø G, Brinchmann-Hansen Å, Hansen KS. Mannequin or standardized patient: participants' assessment of two training modalities in trauma team simulation. *Scand J Trauma Resusc Emerg Med.* 2009;17(1):1–4.
7. Harris DJ, Bird JM, Smart PA, Wilson MR, Vine SJ. A framework for the testing and validation of simulated environments in experimentation and training. *Front Psychol.* 2020;11:605.
8. Mathew RK, Mushtaq F, Immersive Healthcare Collaboration. Three principles for the progress of immersive technologies in healthcare training and education. *Br Med J Simul Technol Enhanc Learn.* 2021;7(5):459–60.

9. Ruthenbeck GS, Reynolds KJ. Virtual reality for medical training: the state-of-the-art. *J Simul.* 2015;9(1):16–26.
10. Semeraro F, Scapigliati A, Ristagno G, Luciani A, Gandolfi S, Lockey A, et al. Virtual Reality for CPR training: how cool is that? Dedicated to the "next generation." *Resuscitation.* 2017;121:e1–2.
11. Bench S, Winter C, Francis G. Use of a virtual reality device for basic life support training: prototype testing and an exploration of users' views and experience. *Simul Healthc.* 2019;14(5):287–92.
12. Barsom EZ, Duijm R, Dusseljee-Peute L, Landman-van der Boom E, van Lieshout E, Jaspers M, et al. Cardiopulmonary resuscitation training for high school students using an immersive 360-degree virtual reality environment. *Br J Educ Technol.* 2020;51(6):2050–62.
13. Buttussi F, Chittaro L, Valent F. A virtual reality methodology for cardiopulmonary resuscitation training with and without a physical mannequin. *J Biomed Inform.* 2020;111:103590.
14. Gent L, Sarno D, Coppock K, Axelrod DM. Successful virtual reality cardiopulmonary resuscitation training in schools: digitally linking a physical manikin to a virtual lifesaving scenario. *Circulation.* 2019;140(2):A396.
15. Kuyt K, Park SH, Chang TP, Jung T, MacKinnon R. The use of virtual reality and augmented reality to enhance cardio-pulmonary resuscitation: a scoping review. *Adv Simul.* 2021;6(1):1–8.
16. Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev.* 2009;1:CD006575.
17. Selvander M, Åsman P. Virtual reality cataract surgery training: learning curves and concurrent validity. *Acta Ophthalmol (Copenh).* 2012;90(5):412–7.
18. Daher S, Hochreiter J, Norouzi N, Gonzalez L, Bruder G, Welch G. Physical-virtual agents for healthcare simulation. 2018. p. 99–106.
19. Mills B, Dykstra P, Hansen S, Miles A, Rankin T, Hopper L, et al. Virtual reality triage training can provide comparable simulation efficacy for paramedicine students compared to live simulation-based scenarios. *Prehosp Emerg Care.* 2020;24(4):525–36.
20. Jaskiewicz F, Kowalewski D, Starosta K, Cierniak M, Timler D. Chest compressions quality during sudden cardiac arrest scenario performed in virtual reality: a crossover study in a training environment. *Medicine (Baltimore).* 2020;99(48):e23374.
21. Bright E, Vine S, Wilson MR, Masters RS, McGrath JS. Face validity, construct validity and training benefits of a virtual reality TURP simulator. *Int J Surg.* 2012;10(3):163–6.
22. Wood G, Wright DJ, Harris D, Pal A, Franklin ZC, Vine SJ. Testing the construct validity of a soccer-specific virtual reality simulator using novice, academy, and professional soccer players. *Virtual Real.* 2021;25(1):43–51.
23. Gray R. Virtual environments and their role in developing perceptual-cognitive skills in sports. In: *Anticipation and decision making in sport.* Routledge; 2019. p. 342–58.
24. Perfect P, Timson E, White MD, Padfield GD, Erdos R, Gubbels AW. A rating scale for the subjective assessment of simulation fidelity. *Aeronaut J.* 2014;118(1206):953–74.
25. Bracq MS, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. *Simul Healthc.* 2019;14(3):188–94.
26. Pan X, Slater M, Beacco A, Navarro X, Bellido Rivas AI, Swapp D, et al. The responses of medical general practitioners to unreasonable patient demand for antibiotics—a study of medical ethics using immersive virtual reality. *PLoS One.* 2016;11(2):e0146837.
27. Slater M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc B Biol Sci.* 2009;364(1535):3549–57.
28. Harris D, Arthur T, de Burgh T, Duxbury M, Lockett-Kirk R, McBarnett W, et al. Assessing Expertise Using Eye Tracking in a Virtual Reality Flight Simulation. *Int J Aeronaut.* 2023. <https://doi.org/10.1080/24721840.2023.2195428>.
29. Usho M, Catena E, Arman S, Slater M. Using presence questionnaires in reality. *Presence.* 2000;9(5):497–503.
30. Harris D, Wilson M, Vine S. Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX). *Virtual Real.* 2020;24(4):557–66.
31. Wilson M, McGrath J, Vine S, Brewer J, Defriend D, Masters R. Psychomotor control in a virtual laparoscopic surgery training environment: gaze control parameters differentiate novices from experts. *Surg Endosc.* 2010;24(10):2458–64.
32. Bright E, Vine SJ, Dutton T, Wilson MR, McGrath JS. Visual control strategies of surgeons: a novel method of establishing the construct validity of a transurethral resection of the prostate surgical simulator. *J Surg Educ.* 2014;71(3):434–9.
33. Słowiński P, Grindley B, Muncie H, Harris D, Vine S, Wilson M. Assessment of cognitive biases in Augmented Reality: Beyond eye tracking. *psyarxiv.com*; 2022 [cited 2022 Dec 15]. Available from: <https://psyarxiv.com/sjyvw/>.
34. Mogg K, Garner M, Bradley BP. Anxiety and orienting of gaze to angry and fearful faces. *Biol Psychol.* 2007;76(3):163–9.
35. Ford BQ, Tamir M, Brunyé TT, Shirer WR, Mahoney CR, Taylor HA. Keeping your eyes on the prize: Anger and visual attention to threats and rewards. *Psychol Sci.* 2010;21(8):1098–105.
36. Zheng B, Jiang X, Tien G, Meneghetti A, Panton ONM, Atkins MS. Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surg Endosc.* 2012;26(10):2746–50.
37. Wu C, Cha J, Sulek J, Zhou T, Sundaram CP, Wachs J, et al. Eye-tracking metrics predict perceived workload in robotic surgical skills training. *Hum Factors.* 2020;62(8):1365–86.
38. Arthur T, Harris D, Buckingham G, Brosnan M, Wilson M, Williams G, et al. An examination of active inference in autistic adults using immersive virtual reality. *Sci Rep.* 2021;11:20377.
39. Mann DL. Predictive processing in the control of interceptive motor actions. In: Cappuccio ML, editor. *Handbook of Embodied Cognition and Sport Psychology.* Cambridge, Massachusetts: The MIT Press; 2019. p. 651–68.
40. Fischer B, Biscaldi M, Otto P. Saccadic eye movements of dyslexic adult subjects. *Neuropsychologia.* 1993;31(9):887–906.
41. Vossel S, Mathys C, Daunizeau J, Bauer M, Driver J, Friston KJ, et al. Spatial attention, precision, and Bayesian inference: a study of saccadic response speed. *Cereb Cortex.* 2014;24(6):1436–50.
42. Amor TA, Reis SD, Campos D, Herrmann HJ, Andrade JS. Persistence in eye movement during visual search. *Sci Rep.* 2016;6(1):20815.
43. Salvucci DD, Goldberg JH. Identifying fixations and saccades in eye-tracking protocols. In *Palm Beach Gardens, Florida, United States: ACM Press*; 2000. p. 71–8.
44. Krassanakis V, Filippakopoulou V, Nakos B. EyeMMV toolbox: An eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification. *J Eye Mov Res.* 2014;7(1):1–10. <https://doi.org/10.16910/jemr.7.1.1>.
45. Lounis C, Peysakhovich V, Causse M. Visual scanning strategies in the cockpit are modulated by pilots' expertise: a flight simulator study. *PLoS One.* 2021;16(2):e0247061.
46. Barby K, Demi B, editors. Detection of driver visual field narrowing. *IEEE*; 2013. p. 634–9.
47. van Doorn J, van den Bergh D, Böhm U, Dablander F, Derks K, Draws T, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon Bull Rev.* 2021;28(3):813–26.
48. Ricci S, Calandrino A, Borgonovo G, Chirico M, Casadio M. Virtual and augmented reality in basic and advanced life support training. *JMIR Serious Games.* 2022;10(1):e28595.
49. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. *Front Robot AI.* 2016;3:74.
50. Harris DJ, Buckingham G, Wilson MR, Vine SJ. Virtually the same? How impaired sensory information in virtual reality may disrupt vision for action. *Exp Brain Res.* 2019;237(11):2761–6.
51. Wijayaratanam DO, Chua R, Cressman EK. Going offline: differences in the contributions of movement control processes when reaching in a typical versus novel environment. *Exp Brain Res.* 2019;237(6):1431–44.
52. Harris DJ, Hardcastle KJ, Wilson MR, Vine SJ. Assessing the learning and transfer of gaze behaviours in immersive virtual reality. *Virtual Real.* 2021;25(4):961–73.
53. Pottle J. Virtual reality and the transformation of medical education. *Future Healthc J.* 2019;6(3):181.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.