Using Mixed Reality Simulation to Improve Junior Medical Trainees’ Preparedness to Manage High-Acuity Trauma

Kirsty Clarke,1,2 Aws Al-mukhtar,1,2,3 Lina Alim,1,2 Amr Nimer,1,2 Sadie Syed,2,3
Emma Hatfield,2 Sanjeeve Sabharwal,2

ABSTRACT
High-acuity trauma necessitates experienced and rapid intervention to prevent patient harm. However, upskilling junior trainees through hands-on management of real trauma cases is rarely feasible without compromising patient safety. This quality education report sought to investigate whether a simulation course operated via mixed reality (MR) headset devices (Microsoft HoloLens) could enhance the clinical knowledge recall and preparedness to practice of junior trainees with no prior experience managing trauma. The Plan–Do–Study–Act quality improvement method was used to refine six emergency trauma vignettes compatible with an MR teaching platform. Each vignette was curated by a multidisciplinary team of orthopaedic surgeons, clinical fellows and experts in simulation-based medical education. As a baseline assessment, a 2-hour emergency trauma course was delivered using traditional didactic methods to a cohort of pre-registration medical students with no clinical exposure to high-acuity trauma (n=16). Next, we delivered the MR simulation to an equivalent cohort (n=32). Clinical knowledge scores derived from written test papers were recorded for each group during and 2 weeks after each course. Each attendee’s end-of-rotation clinical supervisor appraisal grade was recorded, as determined by a consultant surgeon who supervised participants during a 2-week placement on a major trauma ward. Balancing measures included participant feedback and validated cognitive load questionnaires (National Aeronautics and Space Administration–Task Load Index). Overall, attendees of the MR simulation course achieved and sustained higher clinical knowledge scores and were more likely to receive a positive consultant supervisor appraisal. This project serves as a proof of concept that MR wearable technologies can be used to improve clinical knowledge recall and enhance the preparedness to practice of novice learners with otherwise limited clinical exposure to high-acuity trauma.

INTRODUCTION
Many doctors encounter trauma patients for the first time as postgraduates, where lack of knowledge, experience and clinical supervision are associated with increased patient morbidity and mortality.1,2 However, allowing junior team members to build their experience managing high-acuity, low-frequency trauma remains challenging without compromising patient safety.3 Thus, despite strong evidence suggesting that high-quality patient care is best taught through the immersive experiences encountered on placement, most doctors rely on passive learning through lectures and textbooks to instruct their management of high-acuity trauma.4

Encouragingly, the National Health Service (NHS) guidance on the use of medical devices and digital tools,3 referenced in the NHS Long Term Plan,6 recognises the potential for virtual and augmented reality
wearables, such as the Microsoft HoloLens, in providing novice learners with immersive learning experiences which are high fidelity, reproducible and carry no risk of patient harm.\(^7\) As a teaching modality, wearable mixed reality (MR) devices offer andragogy based on constructivism and experiential learning.\(^9\) For example, using HoloHuman and HoloPatient applications (apps), replicas of life-like real patient scenarios can be presented in visual and auditory outputs within an interactive field, allowing students to engage with the scenario in real time.

**Figure 1** A visualisation of how mixed reality (MR) holograms are viewed using the Microsoft HoloLens.

**Figure 2** An adaption of John Sweller’s Cognitive Load Theory describing the learner’s finite capacity for processing information within working memory. Novel technologies can increase extraneous load leading to an unwanted increase in cognitive workload.
from a first-person perspective (figure 1). By encouraging active participation and facilitating the construction of knowledge through experiential learning, MR teaching methods have been shown to increase the retention of clinically relevant knowledge and help learners apply taught knowledge to real patient scenarios. A known drawback of operating MR wearable technologies in time-sensitive scenarios is a noticeable increase in cognitive load, detracting from effective learning. The Cognitive Load Theory (CLT), proposed by John Sweller, posits that learners possess a finite capacity for processing information within their working memory. Figure 2 illustrates how the extraneous load linked to MR technologies must be managed to limit cognitive overload and aid knowledge retention.

The primary objective of this project was to design a novel simulation course where inexperienced learners could interact with simulated high-acuity trauma scenarios in real time via MR headsets. We anticipated learners would maintain optimal levels of cognitive load, improving their ability to recall, retain and apply the principles of high-acuity trauma management.

METHODS

MR simulation development and evaluation

Our NHS Trust receives 12 cohorts of pre-registration doctors for a 2-week trauma and orthopaedics placement. Didactic training focused on emergency trauma management is offered during this time. In this education project, a novel MR simulation course was designed to replace the existing training. MR trauma simulations consisted of six clinical vignettes. Each vignette was created by a multidisciplinary team composed of a senior clinical lecturer in trauma and orthopaedics, surgical teaching fellows and experts in simulation-based medical education. Session development followed a constructivist paradigm and drew upon the concept of CLT. A summary of the lesson
structure can be found in figure 3. The simulation was delivered by surgical teaching fellows using the HoloHuman app on Microsoft HoloLens devices.

As the MR course was complex, with a non-linear relationship between its elements, we agreed to use the CIPP evaluation model for the course.15

► **Context:** Attendees’ needs were assessed using surveys sent out prior to the course.
► **Input:** Experts in surgical education were consulted. Observations of attendees’ behaviour, verbal/written attendee feedback and costing data were used to generate and critique new course ideas.

![PDSA Cycles Diagram]

**Figure 4** Summary of the interventions made in each PDSA cycle compared with the baseline sample. Note that an additional training session teaching students to use the HoloLens was only provided in the last two PDSA cycles. Each 2-week placement was separated by a 4-week interval and the entire project was completed over the course of 24 weeks. MR, mixed reality; PDSA, Plan–Do–Study–Act.
Process: To elicit information about the programme as actually implemented, we asked stakeholders to reflect on how the course aligned with its planned aims.

Product: Attendees received supervisor appraisals 2 weeks after the course, enabling long-term feedback to be gleaned. Stakeholders met after the course and considered methods for improving effectiveness and sustainability.

Quality improvement methodology

The Plan-Do-Study-Act (PDSA) quality improvement method and Standards for Quality Improvement Reporting Excellence guidelines were used (figure 4). Notably, in PDSA 3, students engaged with the teaching remotely from home. The session lead guided students through scenarios using a microphone built into the headset, and learners interacted with the same scenario simultaneously using a shared interactive interface.

Outcome measures for each PDSA cycle included in-placement and follow-up clinical knowledge scores (CKS) from a formative assessment designed to test knowledge management of emergency trauma patients and end-of-placement consultant supervisor report grades for each student (satisfactory, borderline or unsatisfactory). Balancing measures included student feedback and National Aeronautics and Space Administration-Task Load Index (NASA-TLX) questionnaires (employed only for PDSA cycles 2 and 3).

Statistical analysis

Descriptive statistics, including means, SDs and bar graphs, were used to report CKS, NASA-TLX scores, clinical supervisor reports and student feedback. Basic assumptions, including a significant Shapiro-Wilk test for normality, were satisfied for using the Student’s t-test to detect significant differences in mean test scores between cohorts and determine whether in-placement test scores significantly differed from follow-up test scores. Data analysis was done with SPSS Statistics Software IBM V.27.

RESULTS

Outcome measures

The bar graph in figure 5 shows in-placement and follow-up CKS. Supervising consultant appraisals revealed all students (100%) had satisfactory knowledge and ability. Inter-rater agreement between consultant supervisors was 100%.

Independent t-test results (table 1) show that in-placement test scores were significantly higher for students who received MR teaching compared with those who did not (p≤0.01). Paired t-test results show that in-placement test scores were significantly higher than 2-week follow-up test scores, suggesting lower knowledge retention for all cohorts apart from PDSA cycle 3 (p≤0.01).

Balancing measures

NASA-TLX overall scores were not significantly different between cohorts, remaining low for all cohorts receiving MR teaching (means:SD, 8.7±3.32, maximum score 100). Students in PDSA cycle 2 who received MR headset training did not self-report lower Cognitive Workload (CWL) scores than those who did not receive any additional training in using the MR headset (figure 6). Student end-of-session feedback for PDSA cycles 1, 2 and 3 is shown in figure 7.

---

Table 1  t-test results

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean difference</th>
<th>t-value</th>
<th>Significance (two sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent samples t-test comparing CKS achieved by students receiving MR teaching with those receiving lecture-based teaching methods</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CKS*</td>
<td>13.813</td>
<td>3.56</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paired samples t-test comparing CKS achieved during and 2 weeks after teaching, for all PDSA cycles and the baseline sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEM</td>
</tr>
<tr>
<td>Baseline sample</td>
</tr>
<tr>
<td>PDSA 1</td>
</tr>
<tr>
<td>PDSA 2</td>
</tr>
<tr>
<td>PDSA 3</td>
</tr>
</tbody>
</table>

*Equal variances assumed (Levene's test for equality of variances significant at 0.333).
CKS, clinical knowledge scores; MR, mixed reality; PDSA, Plan-Do-Study-Act.
DISCUSSION

In this quality education report (QER), we set out to determine whether a wearable MR teaching platform could be used to deliver high-acuity trauma simulation, therefore, increasing attendees’ ability to recall, retain and apply clinical knowledge and enhancing their readiness to manage real-life emergency trauma scenarios.

Knowledge scores taken during and 2 weeks after the simulation course were significantly higher for those undertaking the MR course than those receiving standard training (p<0.01), indicating better knowledge recall in those using MR headsets. Similarly, the knowledge scores for those who attended the MR simulation were less likely to fall 2 weeks after the training, indicating improved knowledge retention. These findings align with those of a systematic review conducted by Kyaw et al, reporting that extended reality improves knowledge and skills compared with traditional teaching methods. Moreover, a growing body of evidence suggests that the multisensory and interactive features of learning facilitated by MR applications, as demonstrated by platforms like HoloHuman and HoloPatient, more closely resemble the learning experiences occurring in real-life settings. Indeed, unlike traditional simulations with resus dummies, life-like patient interactions are used, incorporating visual elements like changing facial expressions, body language and tone of voice to generate more robust memory traces.

Supervisor appraisals and attendee feedback, including self-reported preparedness to manage trauma (figure 7), were favourable for those taught using MR headsets, indicating that both the novice learners and consultant trauma surgeons recognised an improvement in real-world proficiency for those who undertook the MR simulation. These results are well supported in the simulation literature, as MR teaching methods create situated learning experiences where learners can practise interpreting complex interconnected information unfolding in a time-sensitive manner. As junior doctors make up the main body of the emergency care workforce in most tertiary centres, improving the rate at which they can learn skills in information management, decision-making and emergency response could have significant implications for waiting times and patient flow. Moreover, the utility of MR technologies in upskilling the most junior trainees without forfeiting patient safety cannot be overlooked.

Additionally, it is promising that students in PDSA cycle 3, who undertook their MR session remotely, still achieved significantly higher test scores than those receiving traditional teaching. This finding highlights the potential of MR teaching to transcend geographical barriers, allowing students from disparate locations, including remote or underserved areas, to access high-quality education. However, as identified by Ikeyama et al, effective post-scenario debriefing is essential, especially in the context of emergency trauma teaching. Debriefs may be challenging to facilitate remotely, as encouraging students to address emotions and reflect on their experiences requires tutors to detect subtle non-verbal cues and communicate in a time-sensitive manner.

The instructional design of our MR teaching session was grounded on the principles of CLT proposed by Sweller et al (figure 1). As using new technologies can increase extrinsic workload, we expected the students’ CWL to be high when undertaking MR teaching and improve after MR headset training. However, students’ CWL, measured using the validated NASA-TLX questionnaire, remained low throughout the project, and MR headset training did not significantly reduce CWL scores. Our findings could be explained by the hypothesis that individuals who grew up surrounded by digital technologies (digital natives) have different cognitive abilities and learning styles compared with those who were born before the digital revolution (digital immigrants). Therefore, it is possible that our cohort of undergraduate students had a technological advantage, which resulted in increased cognitive engagement and improved comprehension, unencumbered by unwanted increases in CWL.

Figure 6  Box and whisker plots showing the NASA-TLX results for PDSA cycles 2 and 3. NASA-TLX, National Aeronautics and Space Administration-Task Load Index; PDSA, Plan–Do–Study–Act.

Figure 7  Bar graph showing the results of student feedback questionnaires for PDSA cycles 1, 2 and 3. CKS, clinical knowledge scores; PDSA, Plan–Do–Study–Act.
teaching. Therefore, no comparison between the CWL linked to MR teaching and traditional methods can be drawn. Furthermore, due to the absence of randomisation, pretest and post-test design, it is possible systematic differences in the knowledge and clinical ability of cohorts attending MR teaching versus traditional teaching confounded our results. Future research employing a randomised study design and longer follow-up intervals between knowledge assessments would be of great value in proving cause and effect between MR simulation and knowledge retention. Similarly, a study comparing high-fidelity simulation delivered in a simulation suite and MR simulation delivered via wearable devices would be of merit.

As MR remains an exciting and new way of teaching, novelty bias may have increased student engagement in MR sessions. Therefore, it remains to be seen whether our findings will be sustained once MR technologies are more widely used. Similarly, the follow-up time used to assess knowledge retention was relatively short; further research assessing the impact of MR simulation on knowledge retention over months to years would be of great value.

In addition, despite the potential of wearable MR devices in providing high-fidelity immersive simulation training, the cost of most MR headsets may be prohibitive for many NHS Trusts and undergraduate training institutions. However, it is important to consider that, compared with simulation courses delivered in simulation suits, MR simulation can be run with fewer supporting staff members. Moreover, as MR headset devices become more available and affordable, the provision of high-fidelity simulation using MR is likely to become increasingly cost-effective.

CONCLUSION
This QER demonstrates that emergency trauma simulation provided via MR wearable devices can improve junior medical trainees’ ability to recall, retain and apply clinical anatomy relevant to high-stakes trauma scenarios.

References
14. Stufflebeam DL. The relevance of the CIPP evaluation model for educational accountability.

Open access
This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Kirsty Clarke http://orcid.org/0000-0002-0229-0158
Aws Al-mukhtar http://orcid.org/0000-0002-8257-3448
Lina Alim http://orcid.org/0000-0003-1124-1056

Data availability statement
Data are available upon reasonable request.

Conflict of interest
None declared.