Effectiveness of Immersive Virtual Reality in Surgical Training—A Randomized Control Trial

Yeshwanth Pulijala, MSc, BDS, Minhua Ma, PhD, MSc,† Matthew Pears, MSc, BSc,‡ David Peebles, PhD, MSc,§ and Ashraf Ayoub, PhD, FDS, RCS, MDS, BDS||*

Purpose: Surgical training methods are evolving with the technological advancements, including the application of virtual reality (VR) and augmented reality. However, 28 to 40% of novice residents are not confident in performing a major surgical procedure. VR surgery, an immersive VR (iVR) experience, was developed using Oculus Rift and Leap Motion devices (Leap Motion, Inc, San Francisco, CA) to address this challenge. Our iVR is a multisensory, holistic surgical training application that demonstrates a maxillofacial surgical technique, the Le Fort I osteotomy. The main objective of the present study was to evaluate the effect of using VR surgery on the self-confidence and knowledge of surgical residents.

Materials and Methods: A multisite, single-blind, parallel, randomized controlled trial (RCT) was performed. The participants were novice surgical residents with limited experience in performing the Le Fort I osteotomy. The primary outcome measures were the self-assessment scores of trainee confidence using a Likert scale and an objective assessment of the cognitive skills. Ninety-five residents from 7 dental schools were included in the RCT. The participants were randomly divided into a study group of 51 residents and a control group of 44. Participants in the study group used the VR surgery application on an Oculus Rift with Leap Motion device. The control group participants used similar content in a standard PowerPoint presentation on a laptop. Repeated measures multivariate analysis of variance was applied to the data to assess the overall effect of the intervention on the confidence of the residents.

Results: The study group participants showed significantly greater perceived self-confidence levels compared with those in the control group ($P = .034$; $\alpha = 0.05$). Novices in the first year of their training showed the greatest improvement in their confidence compared with those in their second and third year.

Conclusions: iVR experiences improve the knowledge and self-confidence of the surgical residents.

J Oral Maxillofac Surg 76:1065-1072, 2018

*Doctoral Researcher, School of Art, Design and Architecture, University of Huddersfield, Huddersfield, United Kingdom.

†Professor, Computer Games Technology, School of Computing and Digital Technologies, Staffordshire University, Stoke-on-Trent, United Kingdom.

‡Doctoral Researcher, Department of Psychology, School of Human and Health Sciences, University of Huddersfield, Huddersfield, United Kingdom.

§Reader in Cognitive Science, Department of Psychology, School of Human and Health Sciences, University of Huddersfield, Huddersfield, United Kingdom.

||Professor, Department of Oral and Maxillofacial Surgery, School of Medicine, Dentistry and Nursing, Glasgow University Dental School, Glasgow, United Kingdom.

Conflict of Interest Disclosures: None of the authors have any relevant financial relationship(s) with a commercial interest.

The QR Fund by Higher Education Funding Council for England partially supported travel expenses for data collection by YP and MM. The funders of the study had no role in the data analysis, data interpretation, or writing of the report. The corresponding author had the final responsibility for the decision to submit for publication.

Address correspondence and reprint requests to Dr Pulijala: School of Art, Design and Architecture, University of Huddersfield, Huddersfield HD1 3DH, UK; e-mail: Yeshwanth.pulijala@hud.ac.uk

Received September 4 2017
Accepted October 3 2017

0278-2391/17/31250-8
https://doi.org/10.1016/j.joms.2017.10.002

Self-confidence is considered one of the most influential motivators and regulators of behavior and predicts the successful performance in people's everyday lives.[1,2] The self-confidence of surgeons also influences their performance, professional satisfaction, and success in the future.[3] In a study assessing the errors committed by junior doctors,[4] the largest cause found for both minor and major errors was "feeling overwhelmed." Despite the recent advances in surgical training methods,[5] 28 to 40% of all novice residents have reported not being confident in performing a major procedure.[6,7] The lack of confidence in novices can lead to unintended mishaps during surgery.

A recent systematic review by Elfenbein[8] highlighted the reduced confidence among surgical residents and explained the need for better objective assessments of this attribute. A validated scale for measuring the self-confidence of residents reported that a trainee's confidence in managing a critical surgical situation increases with more exposure to relevant scenarios.[6,9] This practical learning experience with reflection on one's performance is also vital for continuing professional development.[10,11]

However, the reduction in working hours, increased focus on completing more surgical procedures, and inadequate supervision have compromised training.[12] Furthermore, the lack of expertise of the surgical residents at the early stages of their training leads to errors in the operating room, which compromises patient care.[4,13] In oral and maxillofacial surgery (OMS), educational and assessment tools to improve the confidence of the surgical residents are lacking. Furthermore, questions have been raised debating whether the current training is sufficient.[14] A recent review of the European working time directive showed that the reduction in training hours has had a negative effect on some specialties, including OMS, more than on others.[15]

A novice surgical resident usually acquires the fundamental knowledge of surgery, anatomy, and instruments before operating on patients. After achieving a basic competence in the fundamental skills, the residents must overlearn until they develop complementary skills and perform without fear.[16] However, in overcrowded operating rooms, the residents might not obtain the necessary uninterrupted view of the surgical field and thereby would miss essential elements of a surgical procedure. Therefore, a need exists to reform the current surgical training using novel learning tools. Commercially available immersive technologies, including virtual reality (VR) and augmented reality, might provide an answer for these challenges.[17]

VR SURGERY

VR surgery is a holistic learning application, which provides uninterrupted close-up surgical training experience.[18] We used an Oculus Rift development kit (DK2) virtual reality headset and a Leap Motion controller (Leap Motion, Inc, San Francisco, CA) to demonstrate the Le Fort I maxillary osteotomy. This corrective jaw surgery is a complex procedure, which lacks adequate training tools. Furthermore, the constrained surgical field, which is often covered by the surgeon's hands, makes it difficult for the residents to fully observe and master this procedure. To address these challenges, nontechnical skills, including factual knowledge, cognition, and decision making, were highlighted through an enhanced visual experience. The 3 essential elements of VR surgery are a 360° experience of the operating room, close-up stereoscopic visualization of the procedure, and 3-dimensional (3D) interaction. The 360° video creates a sense of presence[19] in the operating room when watched using an Oculus Rift headset (Fig 1). A computer-generated model of the operating room allows the residents to navigate and interact with 3D models of the patient's data, instruments, and anatomy (Fig 2). The cone beam computed tomography scans of the patient, soft tissue planning data, and a surface scan were used in the application. A quiz scene was added to provide real-time feedback to the users. Although the content in this application was limited to the Le Fort I osteotomy, the design and functionality are scalable to other surgical procedures. VR surgery was evaluated in 2 stages, because it is the first immersive VR (iVR) experience for residents in OMS. In the first phase, expert oral and maxillofacial surgeons tested the iVR for face and content validity. The present report discusses the second stage, which evaluated the effect of VR surgery on the residents' knowledge and confidence using a randomized controlled trial (RCT). The aim of the present study was to test the effect of VR surgery on the perceived self-confidence of the residents.

Materials and Methods

DESIGN OF RCT

We evaluated the efficacy of VR surgery in training novices using a multicenter parallel, single-blind RCT. The null hypothesis of the present study was that no difference would result in the perceived self-confidence after intervention between the study and control groups. The alternative hypothesis was that the self-confidence levels of the study group would be different from that of the control group after the intervention.

The researchers read the Declaration of Helsinki on medical protocol. The purpose of the intervention was

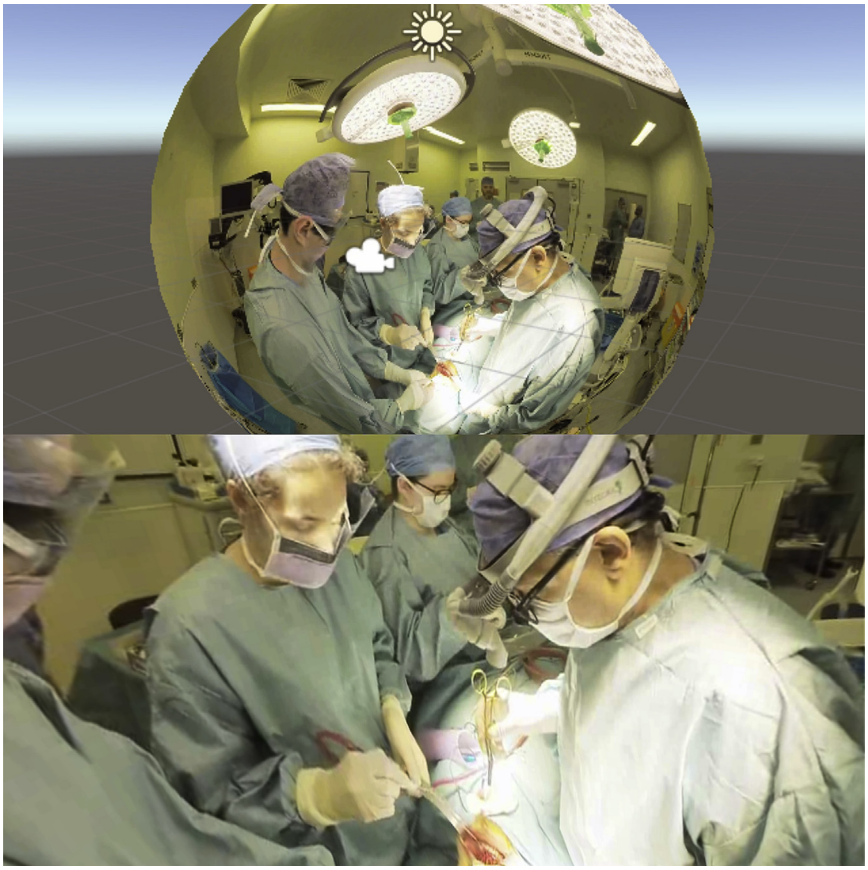

FIGURE 1. Views showing 360° visualization of the operating room in virtual reality surgery.

Pulijala et al. Immersive VR for Surgical Training. J Oral Maxillofac Surg 2018.

to examine the effect on surgical residents only. No patients were involved. The ethics committee of the University of Huddersfield review board approved the present study. All the participants provided written informed consent and participated in the study voluntarily.

OUTCOMES MEASURES

The primary outcomes measure was the comparative evaluation scores of the perceived self-confidence levels before and after the intervention, measured using a 5-point Likert scale. The secondary outcomes were the changes in knowledge levels and the effect of the stage of training on the perceived self-confidence scores.

PARTICIPANT RECRUITMENT

Power calculation using G*Power Analysis[20] for multivariate analysis of variance (ANOVA) revealed the requirement for a sample size of 72 participants for a power of 95 and α value of 0.05. We contacted the head of the OMS departments of 10 dental schools in India and invited their residents to participate in the study. Seven schools responded. After obtaining the necessary permissions, 95 residents were included in the present study. We increased the number of participants to prevent the loss of data through attrition. The study was limited to residents in the full-time master's course of OMS, with limited experience in performing Le Fort I osteotomy. The exclusion criteria were part-time residents who were in their internship, residents with extensive experience in performing the Le Fort I procedure, and participants who could not complete the study.

RANDOMIZATION AND BLINDING

A simple parallel randomization approach was followed in assigning the participants using a randomly generated number series on GraphPad Prism 7 software (GraphPad, San Diego, CA).[21] This, however, resulted in unequal sample size numbers by the end of the study (Fig 3).

STUDY DESIGN

Three questionnaires were designed for the present study. Demographic and preintervention questionnaires were used to provide the baseline data, and the results of the postintervention questionnaire

FIGURE 2. Interaction with 3-dimensional models of the maxillofacial anatomy.

Pulijala et al. Immersive VR for Surgical Training. J Oral Maxillofac Surg 2018.

show the effect of the intervention. Based on previous research on the perceived self-competence by Bandura,[1] a self-confidence scale for the surgical residents in OMS was developed. A questionnaire was designed to accommodate the various elements of confidence needed for a trainee in OMS. A 5-point Likert scale, with 1 indicating the least confident to 5, indicating the most confident, was used to measure this attribute. We queried how the residents perceived their proficiency in the surgical anatomy of the maxilla, instruments used in maxillary osteotomy, and sequence of surgical steps. To counter any inappropriate self-assessment of their confidence,[22] questions testing the knowledge of these aspects were included. To assess the level of situational awareness and decision making, we included 3 questions regarding how the residents would respond to unexpected complications in the operating room and find their weaknesses. To compare the effects of the intervention, we asked these questions before and after the intervention.

Furthermore, we included questions about their learning experience in the operating room and alternative methods of training, including surgical simulators and VR applications. The residents were also able to provide comments on the intervention and give feedback about the best and worst features of the application.

INTERVENTION

The participants required 45 minutes to undergo the intervention. Two supervisors observed the protocol throughout the study period. The study group used VR surgery on an Oculus Rift with Leap Motion tracker, and the control group used a standard power point presentation, which had similar content. For the participants in the study group, the lead researcher demonstrated how to use the system. The residents were asked to interact with the anatomy, data, and instruments routinely used in the surgery through the iVR experience. The participants were asked to watch all the videos clips, including those demonstrating the bone cuts, mobilization of the maxilla, and final fixation of the osteotomy segment. For the control group, stereoscopic 3D videos were replaced by 2-dimensional (2D) videos and 2D images of head and neck anatomy were provided. The 360° videos of operating room were shown on a desktop version of a 360° video viewer, with which the trainee could scroll across the scene with the mouse to watch the operating room ambience.

Results

Among all the participants, 4 residents from control group withdrew from the study after answering the preintervention questionnaire to attend emergency cases in the hospital. The responses of these 4 participants were excluded from the analysis. Of the remaining 91 participants, 48 were male residents (50.5%) and 43 were female residents (45.3%), with a mean age of 27.14 years. A Kolmogorov-Smirnov test of normality was applied to the data ($P > .05$). A visual inspection of the corresponding normality Q-Q plots and histograms showed that the participants' responses followed the normal distribution curve for both the control and the study groups. To ensure that the participants in both groups had a similar level of confidence and knowledge before the intervention, an independent samples t test was performed, which

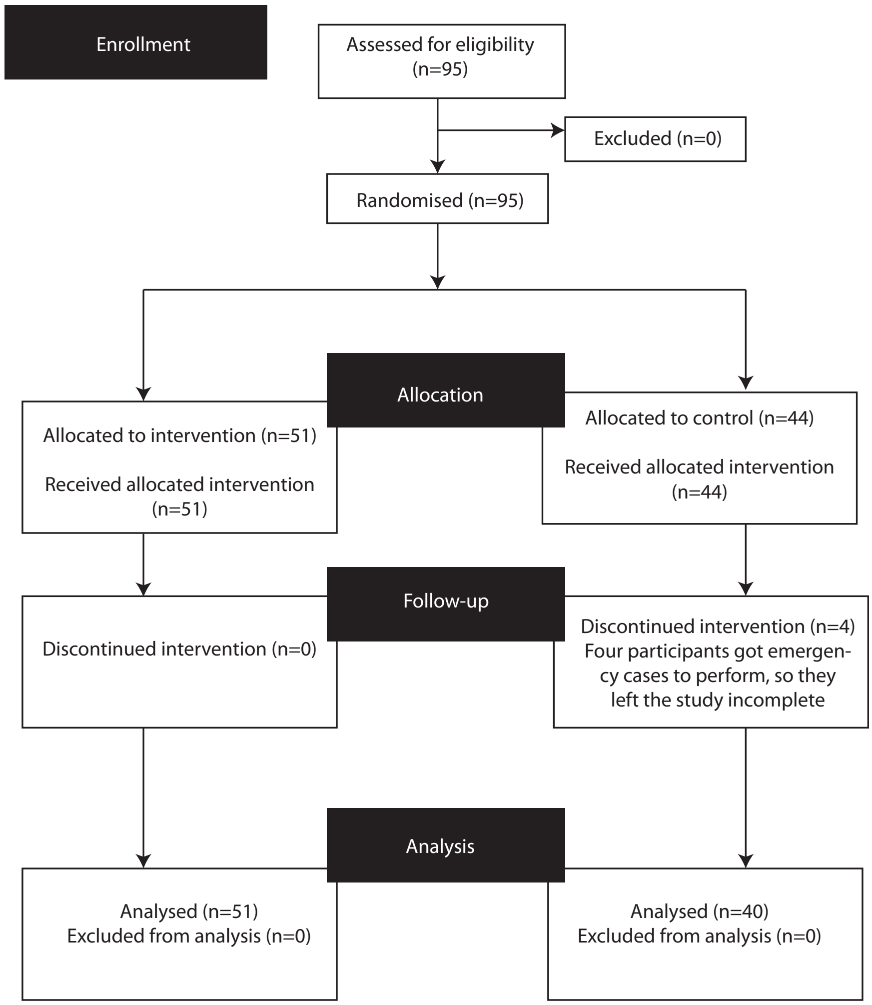

FIGURE 3. Consolidated Standards of Reporting Trials flow diagram for the present randomized control trial.

Pulijala et al. Immersive VR for Surgical Training. J Oral Maxillofac Surg 2018.

showed no significant differences between the 2 groups (t = 0.421; df = 93; *P* = .674).

Repeated measures multivariate ANOVA was applied to the data for the comparative assessment between the overall effect of receiving the VR surgery intervention versus receiving the conventional demonstration on the residents. Although several *t* tests could have been used to compare the responses of the participants in each group, such would have led to many separate *t* tests and increased the risk of a type 1 error.[23] The pre- and postintervention question pairs and intervention groups (study or control) were the within-subject factors. The stage of training was the between-subject factor (Table 1).

Homogeneity of variance assumption using an ANOVA was not violated, as Levene's test showed no statistically significant results. The results showed a significant increase in self-confidence levels [f (1,85) = 65.71; *P* = .000] in both the groups after the intervention. The Wilks lambda multivariate test of the control group showed a statistically significant improvement (*P* = .002) with a small effect size of 0.234 and an observed power of 0.906. In contrast, the participants in the study group showed a statistically significant increase in their confidence (*P* = .000) with a medium effect size of 0.642 and an observed power of 1.000. Comparing the relative improvement in the confidence levels, the participants in the study group showed significantly greater self-confidence scores compared with those in the control group (*P* = .034; Tables 1 and 2); therefore, the null hypothesis was rejected.

The between-subject results showed a significant effect that was dependent on the stage of training [f (2, 85) = 7.57; *P* = .001; partial eta^2 = 0.153; Table 2] of the residents. The post hoc Bonferroni test showed a significant difference between the first-year and third-year residents (*P* = .001); however, the difference between the second-year and third-year residents was not statistically significant (*P* = .360).

Table 1. BETWEEN-SUBJECT FACTORS

Variable	Value Label	Subjects (n)
Group		
1.00	Control	40
2.00	Experimental	51
Stage of study		
1	First-year PG	31
2	Second-year PG	33
3	Third-year PG	27

Abbreviation: PG, postgraduate.

Pulijala et al. Immersive VR for Surgical Training. J Oral Maxillofac Surg 2018.

The VR surgery intervention was found to increase the confidence of early-stage surgical residents.

To assess the effect of the intervention on the knowledge gained, a paired *t* test was performed of each group. The test measured the changes in their mean scores before and after the respective interventions. The paired *t* test results showed a significant increase in the scores for both the control (t = 2.327; df = 43; P = .025) and the study (t = 2.331; df = 50; P = .024) groups. The findings of a 2 (before vs after the intervention) × 2 (experimental vs control group) ANOVA performed to compare the scores of the participants aligned with the nonsignificant improvement in knowledge but a clear pattern of overall improvement. The participants who used the VR surgery performed better than did the control group. When the mean scores of the different questions within the groups were compared, the residents in the study group had a greater mean score for the number of correct answers than did the residents in the control group. They also outperformed the control group for the questions concerning the instruments and sequence of surgical steps. To test the influence of the level of training on the acquired knowledge, we performed a cross-tabs analysis to explore the relationship between the stage of training and the mean score for the correct answers in each group. The results with the greatest improvement were among the first-year surgical residents, followed by the second- and third-year residents in the 2 groups. The difference was more prominent in the study group.

Discussion

Previous studies[3,6] have highlighted a positive correlation between self-confidence and the performance of residents. However, most of the existing studies in OMS did not address issues regarding the self-confidence of the residents. Furthermore, the effect of novel educational interventions such as VR surgery on residents' knowledge and confidence is less known. Our study addressed these questions and our results highlight the future work in surgical training.

At baseline, both the groups had similar scores for self-confidence and knowledge before the intervention. After the intervention, although all the participants had improved their knowledge and confidence, the study group participants outscored the control group. The residents in the study group also showed a significantly greater improvement in their self-confidence after the intervention compared with the participants who had used conventional methods of training. Compared with the control group, the participants in the study group had a

Table 2. MULTIVARIATE TEST* RESULTS

Effect	Value	F	Hypothesis df	Error df	P Value	Partial Eta^2	NC Parameter	Observed Power†
Pre-Post ×								
Pillai's trace	0.436	65.717‡	1.000	85.000	.000	0.436	65.717	1.000
Wilks' lambda	0.564	65.717‡	1.000	85.000	.000	0.436	65.717	1.000
Hotelling's trace	0.773	65.717‡	1.000	85.000	.000	0.436	65.717	1.000
Roy's largest root	0.773	65.717‡	1.000	85.000	.000	0.436	65.717	1.000
Pre-Post group ×								
Pillai's trace	0.052	4.643‡	1.000	85.000	.034	0.052	4.643	0.568
Wilks' lambda	0.948	4.643‡	1.000	85.000	.034	0.052	4.643	0.568
Hotelling's trace	0.055	4.643‡	1.000	85.000	.034	0.052	4.643	0.568
Roy's largest root	0.055	4.643‡	1.000	85.000	.034	0.052	4.643	0.568

Abbreviations: NC, noncomparability; Pre-Post, before to after intervention.

* Design: intercept plus group plus stage of study plus group × stage of study; within-subject design: pair plus Pre-Post plus Pair × Pre-Post.

† Computed using α = 0.05.

‡ Exact statistic.

Pulijala et al. Immersive VR for Surgical Training. J Oral Maxillofac Surg 2018.

compromised learning time because they required some time to become familiar with the technology. Despite these differences, the study group outperformed the control group. This confirms the greater improvement in learning and more comprehensive transfer of knowledge when the residents used the VR surgery application.

The residents credited the holistic experience of the VR surgery for their gain in knowledge and confidence. As justified in previous works,[9] it is logical to assume that with an enhanced knowledge of surgery, anatomy, and instruments, participants will be more confident. Surgical residents greatly appreciated the immersive 360° operating room ambience, 3D interactivity with anatomy and data, and close-up visualization of the surgery, among other features (Video 1).[24] The novel multisensory learning experience might have caused the residents in the study group to experience more confidence than their peers. We noted that 96% of all the participants in the present study had not previously used a virtual reality headset. Hence, the participants who used VR surgery might have experienced a novelty bias that resulted in greater confidence.

In line with previous studies,[6] the stage of training did not have an overall influence on the self-reported confidence levels. However, the post hoc studies revealed that the first-year residents reported significantly greater improvement in their confidence levels compared with that of the second- and third-year residents. The residents in the first year of the training had not observed as many procedures as had the second- and third-year residents. This lack of experience in the operating room might have been the reason the first-year residents showed the most significant improvement in their confidence compared with the other residents.

Improvement in self-confidence is vital for novices in their early stages of training to help them to react appropriately in stressful circumstances. However, a person's perceived self-confidence can also be subject to the Dunning-Kruger effect, a condition in which the ignorant overestimate their ability and performance.[25] To prevent this, we included questions regarding factual knowledge on different aspects of surgery, potential complications, and decision-making skills. The overconfidence of residents should also be monitored and corrected under the supervision of expert surgeons.

Further research should involve a larger sample size to identify the effect of the individual elements of iVR experience on various aspects, including expertise, gender, and the ability to interact. Moreover, because the participants showed a tendency to report an improved sense of confidence immediately after the intervention, it is necessary to test the retention of knowledge and determine whether the levels of self-confidence are maintained for a longer period. Given the differences in the length of OMS training worldwide, it is also desirable to consider a different study population to identify which aspects of VR surgery will be more beneficial for training.

The effect of the attributes acquired with the use of iVR on performance in the operating room also requires investigation. It is not doubted that the application of haptic technology "force feedback" will be an effective addition to iVR for surgical training. As commercially available VR and augmented reality experiences are increasingly used for surgical training,[26] a framework to build effective iVR solutions is needed. We have attempted to address that challenge through a 3-step process of codevelopment, iteration, and evaluation among surgical residents. Currently, the head-mounted VR devices are expensive and require computers with high specifications for a satisfactory VR experience. However, these computers are not easily available at university teaching hospitals and National Health Service institutions.[27] To ensure the global application of these emerging technologies, they must be more affordable. Once the challenges have been met, VR surgery will provide an alternative method of learning that can reduce the time required to train surgeons in the operating room.[28]

Acknowledgments

The authors are indebted to the participants in this study. The authors are thankful to Edward Miller, a virtual reality consultant, who helped in recording the 360° videos for the VR Surgery project.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.joms.2017.10.002.

References

1. Bandura A: Guide for constructing self-efficacy scales, *in* Pajares F, Urdan T (eds): Self-Efficacy Beliefs of Adolescents: Adolescence and Education Series, Vol 5. Charlotte, NC, Information Age Publishing, 2006
2. Cervone D: Thinking about self-efficacy. Behav Modif 24:30, 2000
3. Bucholz EM, Sue GR, Yeo H, et al: Our trainees' confidence: Results from a national survey of 4136 US general surgery residents. Arch Surg 146:907, 2011
4. Baldwin P, Dodd M, Wrate R: Junior doctors making mistakes. Lancet 351:804, 1998
5. Pelargos PE, Nagasawa DT, Lagman C, et al: Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. J Clin Neurosci 35:1, 2017
6. Geoffrion R, Lee T, Singer J: Validating a self-confidence scale for surgical trainees. J Obstet Gynaecol Can 35:355, 2013
7. Rodriguez-Paz J, Kennedy M, Salas E, et al: Beyond "see one, do one, teach one": Toward a different training paradigm. Qual Saf Health Care 18:63, 2009
8. Elfenbein DM: Confidence crisis among general surgery residents: A systematic review and qualitative discourse analysis. JAMA Surg 151:1166, 2016

9. Binenbaum G, Musick DW, Ross HM: The development of physician confidence during surgical and medical internship. Am J Surg 193:79, 2007
10. Flin R, Mitchell L (eds): Safer Surgery: Analysing Behaviour in the Operating Theatre. Boca Raton, FL, CRC Press, 2009
11. Klingensmith ME, Brunt LM: Focused surgical skills training for senior medical students and interns. Surg Clin North Am 90:505, 2010
12. Royal College of Surgeons of England. Taskforce report on the impact of the European Working Time Directive. The Royal College of Surgeons of England. Available at: http://www.rcseng.ac.uk/news-and-events/media-centre/press-releases/taskforce-report-on-impact-of-european-working-time-directive/. Accessed March 2, 2017
13. Elsey EJ, Griffiths G, Humes DJ, West J: Meta-analysis of operative experiences of general surgery trainees during training. Br J Surg 104:22, 2017
14. Hupp JR: Surgical training: Is dabbling enough? J Oral Maxillofac Surg 69:1535, 2011
15. Leahy P: We'll need more than 48 hours to change EWTD rules. RCS Bull 99:41, 2017
16. Verrier ED: The elite athlete, the master surgeon. J Am Coll Surg 224:225, 2017
17. Roy E, Bakr MM, George R: The need for virtual reality simulators in dental education: A review. Saudi Dent J 29:41, 2017
18. Pulijala Y: VR surgery [video]; Vimeo; 2017 Available at: https://vimeo.com/197772557. Accessed March 15, 2017
19. Lombard M, Ditton T: At the heart of it all: The concept of presence. J Comput Commun 3:0, 1997
20. Faul F, Erdfelder E, Buchner A, Lang A-G: Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behav Res Methods 41:1149, 2009
21. GraphPad. GraphPad Prism 7 2017. Available at: https://www.graphpad.com. Accessed February 12, 2017
22. Dunning D, Heath C, Suls JM: Flawed self-assessment implications for health, education, and the workplace. Psychol Sci Publ Interest 5:69, 2004
23. Coolican H: Research Methods and Statistics in Psychology (ed 6). Hove, UK, Psychology Press, 2014
24. Pulijala Y: Surgical trainees' response to virtual reality 2017. Available at: www.youtube.com/watch?v=Cb6yRK4o9mk&feature=youtu.be. Accessed May 4, 2017
25. Dunning D, Johnson K, Ehrlinger J, Kruger J: Why people fail to recognize their own incompetence. Curr Dir Psychol Sci 12:83, 2003
26. Khor WS, Baker B, Amin K, et al: Augmented and virtual reality in surgery—The digital surgical environment: Applications, limitations and legal pitfalls. Ann Transl Med 4:454, 2016
27. Serjeant J: The NHS technology lag cannot go on. 2016 (updated April 17, 2016. Available at: http://www.nuffieldtrust.org.uk/blog/nhs-technology-lag-cannot-go. Accessed April 17, 2016
28. Vinden C, Malthaner R, McGee J, et al: Teaching surgery takes time: The impact of surgical education on time in the operating room. Can J Surg 59:87, 2016