TOPICS IN TRAINING

Improved Complex Skill Acquisition by Immersive Virtual Reality Training

A Randomized Controlled Trial

Ryan Lohre, MD, Aaron J. Bois, MD, MSc, FRCSC, George S. Athwal, MD, FRCSC, and Danny P. Goel, MD, MSc, FRCSC, on behalf of the Canadian Shoulder and Elbow Society (CSES)*

Investigation performed at the Advanced Technical Skills and Simulation Lab (ATSSL), Health Research Innovation Centre, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

Background: There has been limited literature on immersive virtual reality (VR) simulation in orthopaedic education. The purpose of this multicenter, blinded, randomized controlled trial was to determine the validity and efficacy of immersive VR training in orthopaedic resident education.

Methods: Nineteen senior orthopaedic residents (resident group) and 7 consultant shoulder arthroplasty surgeons (expert group) participated in the trial comparing immersive VR with traditional learning using a technical journal article as a control. The examined task focused on achieving optimal glenoid exposure. Participants completed demographic questionnaires, knowledge tests, and a glenoid exposure on fresh-frozen cadavers while being examined by blinded shoulder arthroplasty surgeons. Training superiority was determined by the outcome measures of the Objective Structured Assessment of Technical Skills (OSATS) score, a developed laboratory metric, verbal answers, and time to task completion.

Results: Immersive VR had greater realism and was superior in teaching glenoid exposure than the control ($p = 0.01$). The expert group outperformed the resident group on knowledge testing ($p = 0.04$). The immersive VR group completed the learning activity and knowledge tests significantly faster ($p < 0.001$) at a mean time (and standard deviation) of 11 ± 3 minutes than the control group at 20 ± 4 minutes, performing 3 to 5 VR repeats for a reduction in learning time of 570%. The immersive VR group completed the glenoid exposure significantly faster ($p = 0.04$) at a mean time of 14 ± 7 minutes than the control group at 21 ± 6 minutes, with superior OSATS instrument handling scores ($p = 0.03$). The immersive VR group scored equivalently in surprise verbal scores ($p = 0.85$) and written knowledge scores ($p = 1.0$).

Conclusions: Immersive VR demonstrated substantially improved translational technical and nontechnical skills acquisition over traditional learning in senior orthopaedic residents. Additionally, the results demonstrate the face, content, construct, and transfer validity for immersive VR.

continued

*A list of the Canadian Shoulder and Elbow Society (CSES) members is included in a Note at the end of the article.

Disclosure: The study was supported by the Canadian Shoulder and Elbow Society (CSES). On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work; "yes" to indicate that the author had a patent and/or copyright, planned, pending, or issued, directly relevant to this work (including a pending patent for the immersive virtual reality training module used in the study); and "yes" to indicate that the author had other relationships or activities that could be perceived to influence, or have the potential to influence, what was written in this work (including relationships with PrecisionOS Technologies, the developer of the immersive virtual reality training module used in the study) (http://links.lww.com/JBJS/F681).

Clinical Relevance: This adequately powered, randomized controlled trial demonstrated how an immersive VR system can efficiently (570%) teach a complex surgical procedure and also demonstrate improved translational skill and knowledge acquisition when compared with a traditional learning method.

The Accreditation Council for Graduate Medical Education (ACGME) in North America and the European Working Time Directive (EWTD) promote reduced residency work hours and competency-based education[1]. The FIRST (Flexibility In duty hour Requirements for Surgical Trainees) trial and subsequent publications, comparing flexible work hours with standard ACGME duty-hour requirements in surgical and medical education, have demonstrated that learners can acquire knowledge as measured by passing examinations despite more flexible schedules[2-5]. Unfortunately, these trials do not address technical skill acquisition. Surgical trainees have demonstrated the lack of perceived confidence in independent practice following graduation[6]. Health-care provider expectations are forcing surgical residents to reach competency in the same time as their predecessors[7-9]. To mitigate this, surgical simulation is endorsed by surgical societies including the American Academy of Orthopaedic Surgeons (AAOS) and the American College of Surgeons (ACS)[10,11].

Studies involving surgical simulators receive the Level-of-Evidence statements from the modified Oxford Centre for Evidence-Based Medicine (OCEBM) guidelines[12]. Based on this framework, Level-I evidence of the validity of contemporary immersive virtual reality (VR) for technical and nontechnical skill acquisition in surgical procedures has not been conclusively demonstrated. Immersive VR involves a head-mounted display with visual and auditory cues and controllers using haptic (sense-of-touch) feedback in a 3-dimensional environment. Immersive VR provides realistic operative scenarios with the uninterrupted availability to practice complex surgical skills, devoid of patient-related ethical considerations, patient safety factors, or financial and operative time constraints.

Traditional learning methods include reading educational materials with exposure to procedures, logging case volumes, watching technique videos, performing cadaveric dissections, and receiving feedback from senior colleagues[13]. We tested our hypothesis that immersive VR is superior in teaching a multistep orthopaedic surgical procedure to senior orthopaedic residents compared with traditional learning methods in the form of a journal article. Additionally, we sought to determine the validity, reliability, and transferability of surgical skills obtained in immersive VR training compared with a traditional learning method.

Materials and Methods

Participants

Nineteen orthopaedic surgical residents (resident group) from postgraduate years (PGYs) 4 and 5, and 7 orthopaedic shoulder arthroplasty surgeons (expert group), affiliated with academic institutions from across Canada consented to participate in this study following institutional ethics review board approval. Subjects were recruited from the Canadian Shoulder and Elbow Society (CSES) Annual Resident and Fellow's Course in Calgary, Alberta, Canada. Expert surgeons were defined as those who were fellowship-trained in shoulder surgery, regularly performing arthroplasty and having completed a minimum of 40 shoulder arthroplasties as the primary surgeon.

Randomization

Study participants completed a demographic questionnaire to determine their level of training and experience with simulation. Residents and experts underwent block randomization into equal immersive VR and control groups. The allocated intervention was not revealed to subjects. Figures 1-A and 1-B depict CONSORT (Consolidated Standards of Reporting Trials) flow diagrams for the study.

Intervention

Both resident and expert immersive VR groups performed a glenoid exposure module on an immersive VR platform. This technology utilized a head-mounted display and haptic controllers with an immersive, virtual operating room (Glenoid Exposure Module, version 1.4; PrecisionOS Technologies). The module illustrated key steps in glenoid exposure, including subscapularis takedown, humeral head osteotomy, and soft-tissue releases for difficult glenoid exposure (Figs. 2-A, 2-B, and 2-C). Proper glenoid retractor placement was emphasized. A third-party research member was present during the trial process to mitigate bias. Immersive VR subjects were not limited with respect to time or repetition of the module.

Control subjects were provided with a comprehensive technical journal article outlining steps for achieving glenoid exposure in shoulder arthroplasty similar in educational content to the immersive VR module[14]. Control subjects read the article without time or repetition limitations.

After the learning activity (immersive VR compared with control), both groups completed an 8-question written knowledge test pertaining to exposure, retractor use, and problem-solving for glenoid exposure without referencing the intervention information. All subjects completed a Likert-scale questionnaire on subjective experience with their intervention.

Cadaveric Dissection

Ten additional fellowship-trained shoulder arthroplasty surgeons (evaluators), who were blinded to participant randomization, evaluated residents in a mock operating room scenario. The resident subjects were brought to a surgical skills laboratory in 2 groups of 10 subjects and were instructed to proceed with the necessary surgical steps required to achieve glenoid exposure; all subjects were unaware of the

The Journal of Bone & Joint Surgery · jbjs.org
Volume 102-A · Number 6 · March 18, 2020

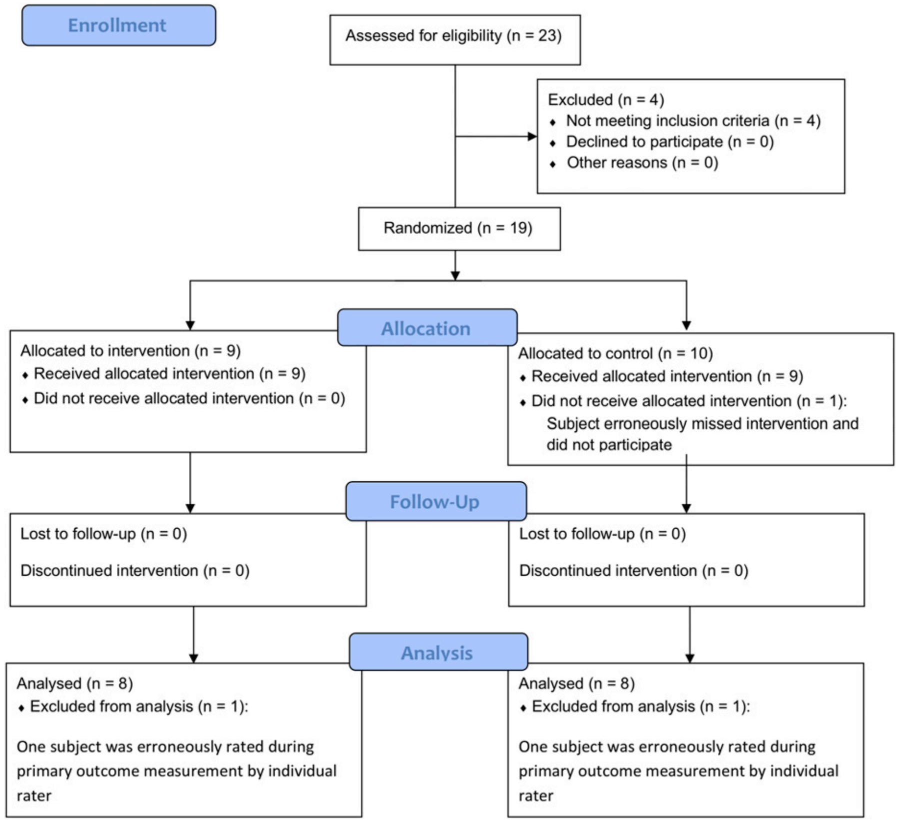

Fig. 1-A

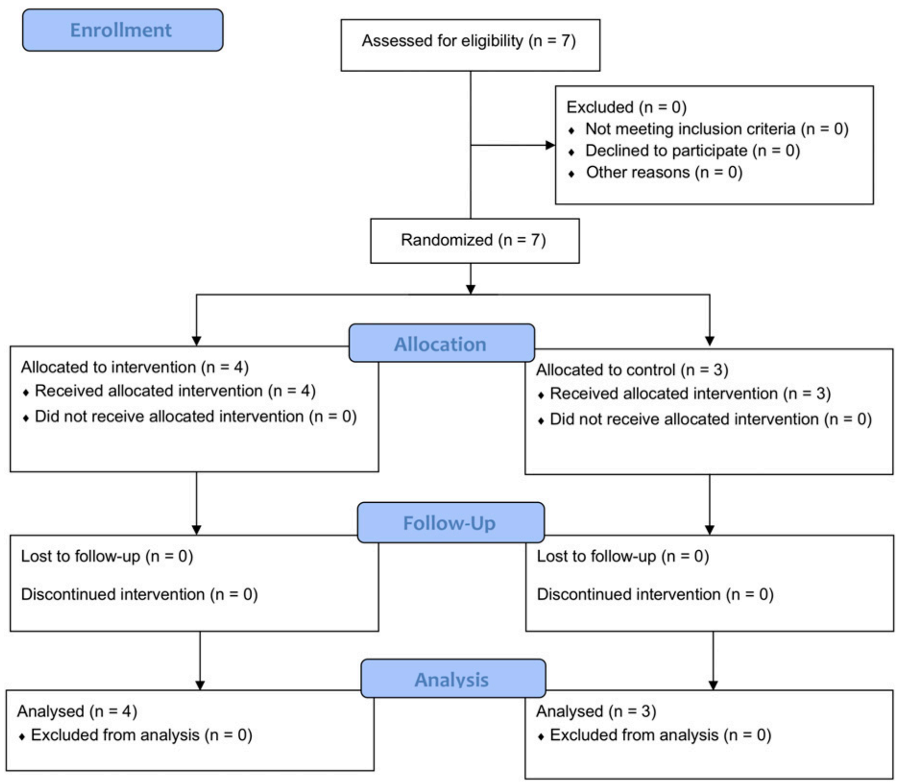

Fig. 1-B

Figs. 1-A and 1-B CONSORT flow diagrams of randomization of resident shoulder surgeons (**Fig. 1-A**) and consultant shoulder surgeons (**Fig. 1-B**).

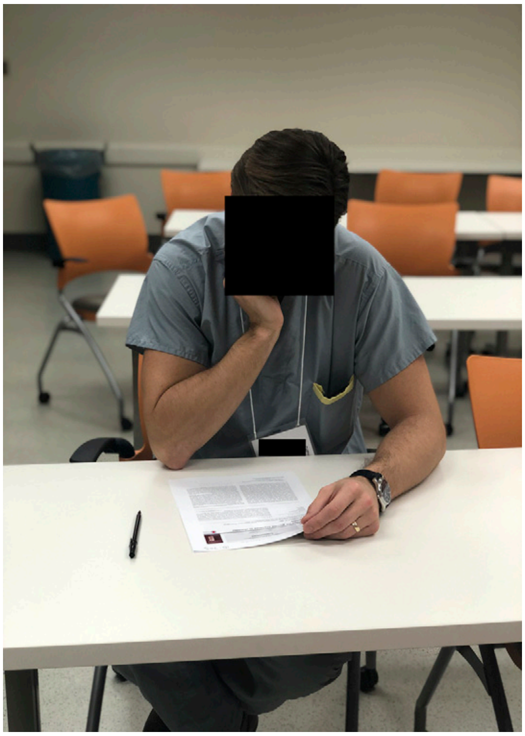

Fig. 2-A

Fig. 2-B

Fig. 2-C

Figs. 2-A, 2-B, and 2-C Representative photographs of learning activities and cadaveric dissection. **Fig. 2-A** A control subject reading the provided technical journal article. **Fig. 2-B** An immersive VR subject performing the virtual glenoid exposure learning activity. **Fig. 2-C** Once the learning activity was completed, the resident group performed a cadaveric dissection and glenoid exposure with an evaluator present.

surgical task being asked of them prior to arrival. The surgical procedure was performed using fresh-frozen upper-extremity cadaveric specimens (from the scapula to the hand). Surgeon evaluators assessed the subjects on the time to completion of glenoid exposure, the Objective Structured Assessment of Technical Skills (OSATS) score, and the completion of a developed laboratory metric[15]. OSATS scoring is a validated metric of open surgery performance[16-18]. The nonvalidated developed

TABLE I Subject Demographic Characteristics

	Resident (N = 16)			Expert (N = 7)			
	Immersive VR (N = 8)	Control (N = 8)	P Value	Immersive VR (N = 4)	Control (N = 3)	P Value	P Value*
Postgraduate training year	4 PYG-4, 4 PGY-5	2 PGY-4, 6 PGY-5	0.63	—	—	—	—
Familiarity with shoulder surgical approaches† ‡	2.22 ± 0.4	2 ± 0	0.14	3 ± 0	3 ± 0	1	<0.001
Familiarity with shoulder arthroplasties† ‡	1.8 ± 0.6	1.8 ± 0.4	0.46	2.5 ± 1	3 ± 0	0.35	<0.001
Shoulder surgical courses attended† ‡	1.1 ± 0.3	1.2 ± 0.4	0.54	3 ± 0	3 ± 0	1	<0.001
Arthroplasties performed as primary surgeon† §	1.2 ± 0.4	1.3 ± 0.5	0.61	3.3 ± 1.5	3.7 ± 0.6	0.35	<0.001
Arthroplasties performed as first assistant† §	2.3 ± 0.5	2 ± 0	0.07	3.3 ± 1.5	3.7 ± 0.6	0.35	<0.001
Yearly shoulder arthroplasties† §	—	—	—	3 ± 1.4	3.3 ± 0.6	0.46	—
Degree of familiarity with surgical simulators† ‡	1 ± 1	1.1 ± 1	0.57	0.5 ± 1	1.3 ± 1.2	0.27	0.65
Use of VR specifically for surgical training#	0	0	1	0	0	1	1
Use of VR in any means prior to study† **	1.1 ± 0.3	1.2 ± 0.4	0.53	1 ± 0	1.7 ± 0.5	0.053	0.51

*Comparison of resident and expert groups. †The values are given as the mean and the standard deviation. ‡Likert scale 1 to 3. §Likert scale 1 to 4. #The values are given as the number of participants. **Likert scale 1 to 2.

laboratory metric pertains to specific sequences of glenoid exposure (see Appendix). Surgeon evaluators were instructed to observe and not provide technical suggestions. The timing of the cadaver glenoid exposure test began once the subscapularis tendon was exposed and was completed once exposure was achieved with retractors positioned and verbal questions answered (Figs. 2-A, 2-B, and 2-C). Two surprise verbal questions were asked during the glenoid exposure task: (1) naming available retractors for exposure, and (2) identifying methods to aid in difficult glenoid exposure. All subjects completed a questionnaire pertaining to the enjoyment, realism, teaching capacity, and perceived longitudinal benefit in continued learning with the repeated use of their respective learning activity.

Outcomes and Statistical Analysis

The primary outcomes were the OSATS score, knowledge score, and time to completion of the cadaveric task. The

TABLE II Primary Outcome Variables

	Novice (N = 16)			Expert (N = 7)			
Variable	Immersive VR* (N = 8)	Control* (N = 8)	P Value	Immersive VR* (N = 4)	Control* (N = 3)	P Value	P Value†
Intervention component							
Knowledge test composite	13.5 ± 1.4	13.5 ± 1.8	1	14.5 ± 3.3	16.7 ± 1.2	0.34	0.04
Time to immersive VR or journal module completion *(min)*	11.0 ± 3.0	20.0 ± 4.0	<0.001	4.0 ± 0	19.0 ± 5.0	0.002	0.12
Cadaveric glenoid exposure component							
OSATS composite	11.8 ± 2.5	12.5 ± 4.8	0.70	—	—	—	—
OSATS: respect for tissue‡	3.0 ± 1.5	3.75 ± 1.0	0.31	—	—	—	—
OSATS: time and motion‡	2.5 ± 0.9	2.8 ± 1.7	0.20	—	—	—	—
OSATS: instrument handling‡	3.25 ± 0.7	3.0 ± 1.8	0.03	—	—	—	—
OSATS: flow of operation and forward planning‡	3.0 ± 1.1	3.0 ± 1.1	1	—	—	—	—
Laboratory metric composite	46.7 ± 5.5	46.2 ± 8.2	0.44	—	—	—	—
Time to glenoid exposure *(min)*	14.0 ± 7.0	21.0 ± 6.0	—	—	—	0.04	—

*The values are given as the mean and the standard deviation. †Comparison of resident and expert groups. ‡Likert scale 1, 3, and 5.

TABLE III Face and Content Validity

	Immersive VR*		Control*		
	Resident (N = 8)	Expert (N = 4)	Resident (N = 8)	Expert (N = 3)	P Value†
Face validity‡					
Overall realism	3.8 (2 to 5)	4 (4 to 4)	2.5 (1 to 4)	3.3 (3 to 4)	<0.001
Equipment realism	4.3 (3 to 5)	4.5 (4 to 5)	—	—	
Anatomy realism	3.9 (3 to 4)	4.25 (4 to 5)	—	—	
Interaction with anatomy realism	3.6 (2 to 4)	4 (4 to 4)	—	—	
Equipment interaction with anatomy realism	3.8 (2 to 4)	4 (4 to 4)	—	—	
Control realism	4 (2 to 5)	3.5 (3 to 4)	—	—	
Instrument movement realism	3.6 (2 to 4)	3.5 (2 to 5)	—	—	
Preparation for operating room activity realism	3.3 (2 to 5)	—	2.7 (2 to 4)	2.6 (1 to 4)	0.34
Overall similarity to operating room	4.1 (2 to 5)	4 (3 to 5)	3.5 (2 to 4)	3 (1 to 4)	0.32
Content validity					
Ability to teach glenoid preparation‡	4.3 (3 to 5)	4.3 (4 to 5)	2.8 (1 to 4)	3.3 (2 to 4)	0.01
Anatomy teaching‡	3.6 (2 to 5)	3.5 (1 to 4)	2.7 (2 to 4)	3.3 (2 to 4)	0.06
Teaching retractor placement‡	4.6 (3 to 5)	4 (3 to 5)	3.3 (2 to 4)	3.6 (3 to 4)	0.005
Teaching problem-solving in glenoid exposure‡	4.1 (3 to 5)	4 (3 to 5)	3.7 (1 to 4)	2.3 (1 to 4)	0.29
Role in surgical education§	3 (3 to 3)	3 (3 to 3)	2.7 (2 to 3)	3 (3 to 3)	0.17

*The values are given as the mean, with the range in parentheses. †Comparison of immersive VR and control in resident group. ‡Likert scale 1 to 5. §Likert scale 1 to 3.

secondary outcome measures included the determination of face, content, construct, and transfer validity. Face validity was defined as the realism of the interventions and content validity was defined as the ability of the interventions to teach the defined task of glenoid exposure and retractor placement, both determined through questionnaires. Construct validity was determined by the comparison of knowledge test scores between expert and resident groups. Transfer validity was demonstration of technical and nontechnical skill transfer from the intervention to the cadaveric dissection, measured as time to completion of the intervention and cadaveric dissection, knowledge, and OSATS scores during dissection.

To achieve 80% statistical power (beta = 0.02) using a 2-sided test at alpha = 0.05, a minimum of 6 subjects were required for each group based on a conservative estimate of 25% difference in knowledge outcome scores of resident and expert groupings. Data were tested for normality prior to statistical analysis. The Student t test was performed for the direct comparison of means for normally distributed data for summative scores and Likert scales. Chi-square testing was performed for normally distributed single Likert-type data. The Cronbach alpha was utilized to determine the reliability of testing metrics and Likert scales. Results were considered significant at $p < 0.05$. Data were handled as a complete case analysis.

Results

Demographic Characteristics

A total of 26 subjects were enrolled in the study. Nineteen orthopaedic surgery residents and 7 expert surgeons were randomized to either the immersive VR group or the control group. One PGY-4 resident (control) was excluded because of missing nearly all portions of the experiment. One PGY-4 resident (control) and 1 PGY-5 resident (immersive VR) were excluded because of an erroneous rating instruction by a single evaluator. The demographic results are summarized in Table I.

There was no difference in resident pre-surgical training between the immersive VR group and the control group ($p = 0.61$). Both cohorts were not familiar with surgical simulation ($p = 0.57$), and none had used VR simulators in previous training ($p = 1.0$) (Table I).

Primary Outcomes

The immersive VR group completed the cadaveric glenoid exposure task faster ($p = 0.04$) at a mean time (and standard deviation) of 14 ± 7 minutes than the control group at 21 ± 6 minutes. The immersive VR group demonstrated superior OSATS instrument handling scores (mean, 3.25 [range, 3 to 5]) compared with the control group (mean, 3.0 [range, 1 to 5]) ($p = 0.03$). There was no difference in the written knowledge scores between the immersive VR group and the control group ($p = 1.0$). The primary outcome results are summarized in Table II.

Secondary Outcomes

Face Validity

The immersive VR activity overall was gauged as realistic by both the resident and expert cohorts. The anatomic features and instrumentation were deemed realistic by both resident and expert cohorts. The haptic feedback demonstrated the lowest

TABLE IV Perceived Benefit of Continued Use

	Resident			Expert		
	Immersive VR* (N = 8)	Control* (N = 8)	P Value	Immersive VR* (N = 4)	Control* (N = 3)	P Value
Enjoyment†						
Enjoyment	4.8 (4 to 5)	3.3 (3 to 4)	<0.001	4.5 (4 to 5)	3.3 (3 to 4)	0.04
Ease of use	4.8 (4 to 5)	4.2 (3 to 5)	0.08	4 (4 to 4)	4.3 (4 to 5)	0.28
Benefit						
Perceived benefit of continued use†	4.5 (4 to 5)	3.2 (2 to 5)	0.009	4.3 (3 to 5)	2.6 (2 to 4)	0.10
Perceived benefit of continued use for novice surgeons‡	3 (3 to 3)	2.6 (2 to 3)	0.08	2.8 (2 to 3)	2.3 (2 to 3)	0.54
Perceived benefit of continued use for expert surgeons‡	2.1 (2 to 3)	2.2 (1 to 3)	0.76	2 (1 to 3)	2 (1 to 3)	1
Perceived benefit to role of immersive VR in surgical education‡	3 (3 to 3)	2.6 (2 to 3)	0.10	3 (3 to 3)	2.6 (2 to 3)	0.29

*The values are given as the mean, with the range in parentheses. †Likert scale 1 to 5. ‡Likert scale 1 to 3.

ratings of realism from both the novice and expert cohorts. Questions of realism had good reliability for both the immersive VR group at 0.82 and the control group at 0.84, as measured by the Cronbach alpha. Overall, the immersive VR group had greater realism (mean, 3.8 [range, 3 to 5]) compared with the control group (mean, 2.5 [range, 1 to 4]) ($p < 0.001$). The details of face validity are summarized in Table III.

Content Validity

Both resident and expert cohorts believed that the immersive VR construct was proficient in teaching. Specifically, resident and expert cohorts believed that the immersive VR construct was, in statistical terms, good with respect to teaching anatomy, glenoid preparation, and glenoid exposure problem-solving and extremely good with respect to teaching retractor placement.

Teaching ability questions had good reliability for the immersive VR group at 0.77 and the control group at 0.74. Immersive VR revealed a greater perceived ability to teach retractor placement than the control module in the resident cohort ($p = 0.005$). The expert cohort similarly believed that the immersive VR module was proficient at teaching glenoid retractor placement and in glenoid exposure and preparation, although it was not significantly different from the control module ($p = 0.21$). The content validity is summarized in Table III.

Construct Validity

On the knowledge test, the expert cohort (15.4 ± 2.7) outperformed the resident cohort (13.6 ± 1.6) ($p = 0.04$). The expert group was able to appropriately name more available retractors (5.3 ± 1.1) than the resident cohort (4.4 ± 0.8) ($p = 0.03$).

Transfer Validity

Time to completion was significantly faster ($p < 0.001$) in the immersive VR group (11 ± 3 minutes) compared with the control group (20 ± 4 minutes). Immersive VR subjects completed 3 to 5 module repetitions, with 3.5 minutes spent per module, accounting for a mean immersive VR module time of 11 ± 3 minutes. This demonstrated a reduction in time of 570% for the resident cohort and 475% for the expert cohort.

The immersive VR resident group was significantly more competent in instrument handling than the control resident group ($p = 0.03$). There was no significant difference ($p = 0.89$) seen in the overall laboratory metric score between the immersive VR group and the control group. The immersive VR resident cohort was able to complete the cadaveric dissection significantly faster ($p = 0.04$) at 14 ± 7 minutes than the control group at 21 ± 6 minutes, a difference of 150%. The immersive VR group scored equivalently to the control group in surprise verbal questioning ($p = 0.85$).

In statistical terms, the OSATS score demonstrated good reliability (0.78) and the laboratory metric score showed very good reliability (0.84), as measured by the Cronbach alpha for both the immersive VR and control groups.

Further Benefit Consideration

The resident immersive VR group enjoyed the learning activity (mean, 4.8 [range, 4 to 5]) more than the resident control group (mean, 3.3 [range, 3 to 4]) ($p < 0.001$). The resident immersive VR group perceived an educational benefit to continued use, an educational benefit to novice and expert surgeons, and an overall benefit to surgical education. The expert immersive VR cohort similarly enjoyed the immersive VR activity (mean, 4.5 [range, 4 to 5]) more than the control group enjoyed the article reading (mean, 3.3 [range, 3 to 4]) ($p = 0.04$). Additionally, the expert immersive VR cohort believed that there was a benefit with repeated use for novice surgeons and surgical education. Table IV demonstrates Likert responses to the perceived benefit of continued use.

Discussion

In 1993, Satava proposed VR surgical simulation as the future of surgical education[19]. Numerous surgical simulators have subsequently been produced to provide technical and nontechnical skills[1,20-28]. Previously, VR was defined with simulators that were largely single-use, were of varying fidelity, and had limited validity and reliability despite direct recommendations by the Work Group for Evaluation and Implementation of Simulators and Skills Training Programmes of the European Association of Endoscopic Surgeons (EAES)[12]. The recent Innovation, Design, and Emerging Alliances in Surgery (IDEAS) Conference identified the production of an immersive environment with an educational focus on skill acquisition, adaptive learner performance, and surgical procedure rehearsal as key parameters to a successful implementation of a VR training program[29]. To our knowledge, our study is the first to achieve these goals and demonstrate the both technical and nontechnical skill transfer in surgical trainees by contemporary immersive VR training.

By means of a blinded, randomized controlled trial, an immersive VR-trained group of senior orthopaedic surgical residents demonstrated superior instrument handling by the validated OSATS score. The immersive VR group was able to perform a complex, multistep surgical task following an instructional session faster than a traditionally trained group, with equivalent retained knowledge scores tested through written and oral testing. Block randomization produced a similar level of training between the immersive VR group and the control group. The improved instrument handling demonstrated may come from the immersion aspect of the VR module, which allows users to perform, rather than simply outline, key surgical steps. Learning through repetition may also promote the improvement of learning and technical skill acquisition, as the immersive VR group was able to achieve multiple passes of the module compared with the control group. Kohls-Gatzoulis et al. highlighted the importance of multifaceted training for cognitive and technical skill acquisition through junior resident learning of total knee replacements[30]. This work elaborates Ericsson's concept of deliberate practice[31]. Cognitive capacity and decision-making can be taught through repetition and are important skills for any surgeon.

Surgical trainees now have flexible work hours following the FIRST trial[2]. Confidence in learning through this evolution has been questioned, with surgical simulators proposed as a means to compensate for the heterogeneity of case experiences[5,6]. Despite the reduction in work hours, concern about burnout remains in the forefront of surgical education, with estimates of 27% to 75% of residents experiencing evidence of burnout[32]. A previous study of orthopaedic surgery trainees showed a clear dependency of overall satisfaction on reduced working time[33]. Unspecified in these studies is the amount of learning time in addition to direct work hours, which can be overwhelming. With the tasks examined, we have demonstrated significant improvements of technical and nontechnical skill acquisition at a mean reduction of 570% in learning time, with the ability to repeat complex multistep tasks while being substantially more enjoyable than the traditional learning method.

Surgical simulation has been identified as a means to improve the quality of care in low and middle-income populations in general surgery[34]. Fractures incur 52 million disability-adjusted life years annually around the world and account for 16% of the total burden of disease globally[35]. An inverse relationship exists between the number of surgical procedures performed in low and middle-income countries and the rates of morbidity and mortality[36]. The acquisition of technical and nontechnical skills through immersive VR could dramatically impact the existing health-care disparity noted and could have vast socioeconomic impacts on the working population in these developing countries.

This study has a number of limitations. It had a small sample size, and convenience sampling precluding the inclusion of an intermediate cohort (i.e., fellows or junior consultants). Transfer validity was determined using cadaveric specimens rather than a real surgical scenario; however, cadavers have been recommended as an appropriate adjunct in studies of transfer validity[7-9]. Expert surgeons did not complete the cadaveric dissection because of inherent bias by colleague evaluators and limits on further construct validation of technical skill acquisition. There is not a well-established rating scale of open surgery performance for orthopaedic surgery; however, OSATS scores have been previously validated in the assessment of open shoulder surgical procedures[18,37,38]. The establishment of other validity metrics including face and content validity is very subjective, and there is no well-established questionnaire to address this. We used a nonvalidated laboratory metric for the additional assessment of residents during cadaveric dissection; however, it demonstrated very good reliability in its scoring and association with OSATS scores. The study design could be improved by using multiple evaluators per study participant or recording and analysis by multiple evaluators. Our study exclusively compared immersive VR with a technical journal article; however, several other modalities of traditional training exist including videos, cadaveric dissections, instructional courses, or bench-top simulators. The role of immersive VR in orthopaedic surgery requires further validation in comparison with these training methods before a definitive recommendation of equivalency and generalizability in training can occur. This study exclusively examined 1 modality of training compared with immersive VR in a select population of senior learners and thus has limited generalizability to other levels of training or procedures. VR simulators have previously received poor sense-of-touch ratings, and our haptics demonstrated room for improvement, as shown in Table III. A further limitation is the single use of the intervention rather than longitudinal assessment.

In conclusion, immersive VR was effective in teaching a single complex surgical skill to senior-level residents and was superior to a traditional method of learning in time to task completion and instrument handling in this blinded, randomized controlled trial. The immersive VR learning task provided equivalent nontechnical skill to trainees measured via knowledge tests and was significantly more efficient for resident and expert surgeons than traditional teaching. In this study, immersive VR demonstrated the ability to teach a complex surgical skill efficiently with face, content, construct, and transfer validity. However, further comparative validity studies are required for immersive VR in training orthopaedic surgeons.

Appendix

eA Supporting material provided by the authors is posted with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJS/F682). ■

Note: The members of the Canadian Shoulder and Elbow Society (CSES) include Dr. Ryan Bicknell, Dr. Marty Bouliane, Dr. Patrick Chin, Dr. Jonah Davies, Dr. Darren Drosdowech, Dr. Justin LeBlanc, Dr. Jason Old, Dr. J. Pollock, Dr. Dominique Rouleau, Dr. Marlis Sabo, and Dr. David Sheps.

The authors thank the study participants (and current members of the CSES): Dr. Rob Balyk, Dr. Cole Beavis, Dr. Ken Faber, Dr. Diane Nam, Dr. Peter Lapner, Dr. Dave Saliken, and Dr. William Regan.

Ryan Lohre, MD[1]
Aaron J. Bois, MD, MSc, FRCSC[2]
George S. Athwal, MD, FRCSC[3]
Danny P. Goel, MD, MSc, FRCSC[1]

[1]Department of Orthopaedics, University of British Columbia, Vancouver, British Columbia, Canada

[2]Section of Orthopaedic Surgery, Department of Surgery, University of Calgary, Calgary, Alberta, Canada

[3]Roth McFarlane Hand and Upper Limb Center, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

Email address for D.P. Goel: danny.goel@ubc.ca

ORCID iD for R. Lohre: 0000-0002-3915-888X
ORCID iD for A.J. Bois: 0000-0002-4883-6566
ORCID iD for G.S. Athwal: 0000-0002-5174-219X
ORCID iD for D.P. Goel: 0000-0001-5894-685X

References

1. Stirling ERB, Lewis TL, Ferran NA. Surgical skills simulation in trauma and orthopaedic training. J Orthop Surg Res. 2014 Dec 19;9:126-34.
2. Bilimoria KY, Chung JW, Hedges LV, Dahlke AR, Love R, Cohen ME, Hoyt DB, Yang AD, Tarpley JL, Mellinger JD, Mahvi DM, Kelz RR, Ko CY, Odell DD, Stulberg JJ, Lewis FR. National cluster-randomized trial of duty-hour flexibility in surgical training. N Engl J Med. 2016 Feb 25;374(8):713-27. Epub 2016 Feb 2.
3. Saadat LV, Dahlke AR, Rajaram R, Kreutzer L, Love R, Odell DD, Bilimoria KY, Yang AD. Program director perceptions of surgical resident training and patient care under flexible duty hour requirements. J Am Coll Surg. 2016 Jun;222(6):1098-105. Epub 2016 Apr 4.
4. Desai SV, Asch DA, Bellini LM, Chaiyachati KH, Liu M, Sternberg AL, Tonascia J, Yeager AM, Asch JM, Katz JT, Basner M, Bates DW, Bilimoria KY, Dinges DF, Even-Shoshan O, Shade DM, Silber JH, Small DS, Volpp KG, Shea JA; iCOMPARE Research Group. Education outcomes in a duty-hour flexibility trial in internal medicine. N Engl J Med. 2018 Apr 19;378(16):1494-508. Epub 2018 Mar 20.
5. Blay E Jr, Hewitt DB, Chung JW, Biester T, Fiore JF, Dahlke AR, Quinn CM, Lewis FR, Bilimoria KY. Association between flexible duty hour policies and general surgery resident examination performance: a Flexibility In duty hour Requirements for Surgical Trainees (FIRST) trial analysis. J Am Coll Surg. 2017 Feb;224(2):137-42. Epub 2016 Nov 21.
6. George BC, Bohnen JD, Williams RG, Meyerson SL, Schuller MC, Clark MJ, Meier AH, Torbeck L, Mandell SP, Mullen JT, Smink DS, Scully RE, Chipman JG, Auyang ED, Terhune KP, Wise PE, Choi JN, Foley EF, Dimick JB, Choti MA, Soper NJ, Lillemoe KD, Zwischenberger JB, Dunnington GL, DaRosa DA, Fryer JP; Procedural Learning and Safety Collaborative (PLSC). Readiness of US general surgery residents for independent practice. Ann Surg. 2017 Oct;266(4):582-94.
7. Holland JP, Waugh L, Horgan A, Paleri V, Deehan DJ. Cadaveric hands-on training for surgical specialties: is this back to the future for surgical skills development? J Surg Educ. 2011 Mar-Apr;68(2):110-6.
8. Gilbody J, Prasthofer AW, Ho K, Costa ML. The use and effectiveness of cadaveric workshops in higher surgical training: a systematic review. Ann R Coll Surg Engl. 2011 Jul;93(5):347-52.
9. Akhtar KSN, Chen A, Standfield NJ, Gupte CM. The role of simulation in developing surgical skills. Curr Rev Musculoskelet Med. 2014 Jun;7(2):155-60.
10. Aïm F, Lonjon G, Hannouche D, Nizard R. Effectiveness of virtual reality training in orthopaedic surgery. Arthroscopy. 2016 Jan;32(1):224-32. Epub 2015 Sep 26.
11. Vergis A, Steigerwald S. Skill acquisition, assessment, and simulation in minimal access surgery: an evolution of technical training in surgery. Cureus. 2018 Jul 12;10(7):e2969-2969.
12. Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB, Jakimowicz JJ. Consensus guidelines for validation of virtual reality surgical simulators. Simul Healthc. 2006 Fall;1(3):171-9.
13. Losco M, Familiari F, Giron F, Papalia R. Use and effectiveness of the cadaver-lab in orthopaedic and traumatology education: an Italian survey. Joints. 2017 Dec 11;5(4):197-201.
14. Williams GR Jr. Top ten tips for glenoid exposure in shoulder arthroplasty. Semin Arthroplasty. 2017 Sep;28(3):124-7.
15. Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. Acad Med. 1996 Dec;71(12):1363-5.
16. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997 Feb;84(2):273-8.
17. Van Nortwick SS, Lendvay TS, Jensen AR, Wright AS, Horvath KD, Kim S. Methodologies for establishing validity in surgical simulation studies. Surgery. 2010 May;147(5):622-30. Epub 2009 Dec 16.
18. Bernard JA, Dattilo JR, Srikumaran U, Zikria BA, Jain A, LaPorte DM. Reliability and validity of 3 methods of assessing orthopedic resident skill in shoulder surgery. J Surg Educ. 2016 Nov-Dec;73(6):1020-5. Epub 2016 Jun 3.
19. Satava RM. Virtual reality surgical simulator. The first steps. Surg Endosc. 1993 May-Jun;7(3):203-5.
20. Yari SS, Jandhyala CK, Sharareh B, Athiviraham A, Shybut TB. Efficacy of a virtual arthroscopic simulator for orthopaedic surgery residents by year in training. Orthop J Sports Med. 2018 Nov 21;6(11):2325967118810176.
21. Waterman BR, Martin KD, Cameron KL, Owens BD, Belmont PJJ Jr. Simulation training improves surgical proficiency and safety during diagnostic shoulder arthroscopy performed by residents. Orthopedics. 2016 May 1;39(3):e479-85. Epub 2016 May 2.
22. Morgan M, Aydin A, Salih A, Robati S, Ahmed K. Current status of simulation-based training tools in orthopedic surgery: a systematic review. J Surg Educ. 2017 Jul-Aug;74(4):698-716. Epub 2017 Feb 8.
23. McDougall EM. Validation of surgical simulators. J Endourol. 2007 Mar;21(3):244-7.
24. Li L, Yu F, Shi D, Shi J, Tian Z, Yang J, Wang X, Jiang Q. Application of virtual reality technology in clinical medicine. Am J Transl Res. 2017 Sep 15;9(9):3867-80.
25. Kim Y, Kim H, Kim YO. Virtual reality and augmented reality in plastic surgery: a review. Arch Plast Surg. 2017 May;44(3):179-87. Epub 2017 May 22.
26. Hung AJ, Zehnder P, Patil MB, Cai J, Ng CK, Aron M, Gill IS, Desai MM. Face, content and construct validity of a novel robotic surgery simulator. J Urol. 2011 Sep;186(3):1019-24. Epub 2011 Jul 23.
27. Dorozhkin D, Olasky J, Jones DB, Schwaitzberg SD, Jones SB, Cao CGL, Molina M, Henriques S, Wang J, Flinn J, De S; SAGES FUSE Committee. OR fire virtual training simulator: design and face validity. Surg Endosc. 2017 Sep;31(9):3527-33. Epub 2016 Dec 30.
28. Bartlett JD, Lawrence JE, Stewart ME, Nakano N, Khanduja V. Does virtual reality simulation have a role in training trauma and orthopaedic surgeons? Bone Joint J. 2018 May 1;100-B(5):559-65.
29. Olasky J, Sankaranarayanan G, Seymour NE, Magee JH, Enquobahrie A, Lin MC, Aggarwal R, Brunt LM, Schwaitzberg SD, Cao CG, De S, Jones DB. Identifying opportunities for virtual reality simulation in surgical education: a review of the proceedings from the Innovation, Design, and Emerging Alliances in Surgery (IDEAS) conference: VR surgery. Surg Innov. 2015 Oct;22(5):514-21. Epub 2015 Apr 29.
30. Kohls-Gatzoulis JA, Regehr G, Hutchison C. Teaching cognitive skills improves learning in surgical skills courses: a blinded, prospective, randomized study. Can J Surg. 2004 Aug;47(4):277-83.
31. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med. 2004 Oct;79(10)(Suppl):S70-81.

32. Ishak WW, Lederer S, Mandili C, Nikravesh R, Seligman L, Vasa M, Ogunyemi D, Bernstein CA. Burnout during residency training: a literature review. J Grad Med Educ. 2009 Dec;1(2):236-42.

33. Daniels AH, DePasse JM, Kamal RN. Orthopaedic surgeon burnout: diagnosis, treatment, and prevention. J Am Acad Orthop Surg. 2016 Apr;24(4):213-9.

34. Tansley G, Bailey JG, Gu Y, Murray M, Livingston P, Georges N, Hoogerboord M. Efficacy of surgical simulation training in a low-income country. World J Surg. 2016 Nov;40(11):2643-9.

35. Stewart BT, Gyedu A, Tansley G, Yeboah D, Amponsah-Manu F, Mock C, Labi-Addo W, Quansah R. Orthopaedic trauma care capacity assessment and strategic planning in Ghana: mapping a way forward. J Bone Joint Surg Am. 2016 Dec 7; 98(23):e104-104.

36. Myles PS, Haller G. Global distribution of access to surgical services. Lancet. 2010 Sep 25;376(9746):1027-8. Epub 2010 Jul 2.

37. VanHeest A, Kuzel B, Agel J, Putnam M, Kalliainen L, Fletcher J. Objective structured assessment of technical skill in upper extremity surgery. J Hand Surg Am. 2012 Feb;37(2):332-7: 337.e1-4.

38. Anderson DD, Long S, Thomas GW, Putnam MD, Bechtold JE, Karam MD. Objective structured assessments of technical skills (OSATS) does not assess the quality of the surgical result effectively. Clin Orthop Relat Res. 2016 Apr;474(4):874-81.