Transfer Effectiveness of Surgical Education Using Immersive Virtual Reality: A Randomized Intervention-Controlled Trial

Ryan Lohre, MD
University of British Columbia
Vancouver, British Columbia, Canada
Ryan.lohre@alumni.ubc.ca

Aaron Bois, MD, MSc
University of Calgary
Calgary, Alberta, Canada
ajmbois@gmail.com

George S Athwal, MD
Western University
London, ON, Canada
gsathwal@hotmail.com

J Pollock, MD, MSc, Peter Lapner, MD
University of Ottawa
Ottawa, ON, Canada
jpollock@toh.ca, plapner@me.com

Danny P. Goel, MD, MBA, MSc, FRCSC
University of British Columbia
Vancouver, British Columbia
Danny.goel@ubc.ca

Canadian Shoulder and Elbow Society
4060 St. Catherine Street West, Suite 620
Westmount, Quebec H3Z 2Z3
T: 514 874-9003 F: 514 874-0464
info@canorth.org

ABSTRACT

Background: Immersive virtual reality (iVR) simulators are increasingly used in surgical training. The purpose of this study was to determine the effectiveness of iVR training compared to video training in complex skill acquisition. Our hypothesis was that iVR improved learning effectiveness and a virtual reality ratings scale correlated to real-world performance.

Methods: Eighteen orthopaedic surgery residents (PGY4-5) from multiple institutions were randomized to receive training for reverse shoulder arthroplasty (RSA). The experimental group (n=9) trained with iVR compared to the control group (n=9) with traditional technical instructional video. Residents completed both demographic questionnaires and written/verbal knowledge tests. Each resident was evaluated by a blinded shoulder surgeon on performing an RSA on a fresh-frozen cadaver. Measurements included the Objective Structured Assessment of Technical Skills (OSATS) and Global Ratings Scale (GRS). Resident performance metrics were recorded.

Results: Both groups were demographically similar. The iVR group completed training 155%, 95% CI [1.64,10.3]; $p=0.008$ faster with improved verbal questioning scores ($M = 4.1$, $SD = 1.0$; vs $M = 2.2$, SD 1.7; 95% CI [0.10,3.3]; $p=0.03$). The iVR group ($M = 15.9$, $SD = 2.5$; 95% CI [14.3,17.5]) had significantly better 95% CI [3.3,9.7;$p<0.001$) OSATS scores than the control group ($M = 9.4$, $SD = 3.2$; 95% CI [7.4,11.5]). The iVR training scores had a strong correlation to real-world OSATS scores ($r(8) = 0.74$, $p = 0.035$) and implant position ($r(12) = 0.73$, $p = 0.006$). The transfer of training (ToT) was 32.6-% - 59.4% based on comparative completion time and OSATS scores, respectively. The transfer effectiveness ratio (TER) was 0.79.

Conclusion: Training senior surgical residents with iVR demonstrated superior learning efficiency, knowledge and skill transfer as measured by ToT. Based on published RSA learning curves, the iVR module reduces early training by 13-51 operative cases. The TER of 0.79 is higher than prominent surgical simulators and average flight simulators, providing 47.4 minutes of saved OR time for 60 minutes of iVR training. This study demonstrates that iVR should be considered as an efficient, safe and effective means for procedural skill training.

ABOUT THE AUTHORS

Dr. Ryan Lohre is a post-graduate year 5 (PGY5) orthopaedic surgery resident at the University of British Columba (UBC) in Vancouver, British Columbia, Canada.
Dr. Aaron Bois is a consultant shoulder surgeon and clinical associate professor at the University of Calgary in Calgary, Alberta, Canada.

Dr. George Athwal is a consultant shoulder surgeon at the Roth McFarlane Hand and Upper Limb Clinic and professor in the Schulich School of Medicine and Dentistry at Western University in London, Ontario, Canada.
Dr. J Pollock is a consultant shoulder surgeon and assistant professor at the University of Ottawa Medical school in Ottawa, ON, Canada
Dr. Peter Lapner is a consultant shoulder surgeon and associate professor at the University of Ottawa Medical school in Ottawa, ON, Canada

Canadian Shoulder Elbow Society: 4060 St. Catherine Street West, Suite 620
Westmount, Quebec H3Z 2Z3
T: 514 874-9003 F: 514 874-0464
info@canorth.org

Dr. Danny P Goel is a consultant shoulder surgeon and clinical associate professor with the University of British Columbia Department of Orthopaedics. He is the co-founder and CEO of Precision OS Technology

Transfer Effectiveness of Surgical Education Using Immersive Virtual Reality: A Randomized Intervention-Controlled Trial

Ryan Lohre, MD	Aaron Bois, MD, MSc	George S Athwal, MD	J Pollock, MD, MSc	Peter Lapner, MD	Danny P. Goel, MD, MBA, MSc
University of British Columbia	University of Calgary	Western University	University of Ottawa	University of Ottawa	University of British Columbia
Vancouver, British Columbia, Canada	Calgary, Alberta,	London, ON, Canada	Ottawa, ON, Canada	Ottawa, ON, Canada	Vancouver, British Columbia
Ryan.lohre@alumni.ubc.ca	ajmbois@gmail.com	gsathwal@hotmail.com	jpollock@toh.ca	plapner@me.com	Danny.goel@ubc.ca

INTRODUCTION

Simulator use in procedural education of health care providers is prevalent around the world. The use of simulators is supported by studies examining the comparison of, and transfer of skill training, to the real world (Bjerrum, Konge, Subhi, & Thomsen, 2017). Examining the degree of skill transfer with training time using simulation could help to reduce real-world training time (Korteling, Oprins, & Kallen, 2009) and significantly reduce costs. Simulator analysis and validation has been used extensively in the aviation industry and military complex (Bartlett, Lawrence, Stewart, Nakano, & Khanduja, 2018). This analysis method has been less extensively studied in healthcare.

Combat casualty care is not reflective of civilian training and sub-specialist availability (Breeze, Combes, DuBose, & Powers, n.d.). Military surgical personnel require extended procedural knowledge when compared to their civilian counterparts. The US military provides service specific training through pre-deployment courses and clinical rotations in level 1 trauma centers. Courses provide instruction through traditional classroom and cadaveric experiences.(Breeze et al., n.d.) Cadaveric training is costly, single use, and lacks evidence of skill acquisition and retention. Even in high-volume trauma centers, heterogeneity of experience may predispose to knowledge gaps. There may be a role for expanded procedural training incorporating novel surgical simulators.

Immersive virtual reality (iVR) offers a portable, multisensory, cost effective training experience. The learning environment is consistent and scalable to experience level. Immersive VR has been validated during skills training in orthopedic surgery.(Logishetty, Gofton, Rudran, Beaule, & Cobb, 2020; Lohre, Bois, Athwal, Goel, & (CSES), 2020) Training time using iVR has been shown to be significantly faster and simultaneously more enjoyable than traditional training methods.(Lohre, Bois, et al., 2020) However, further skill transfer must be studied to continue to demonstrate the value of iVR relative to real life experiences. This will additionally inform the cost effectiveness of procedural training using iVR.

Traditional simulators rely on validated scoring metrics to determine effectiveness. Immersive VR should have a similar validated metric given its increasing use in procedural education. An iVR scoring metric could be used to determine longitudinal skill acquisition and retention. It could also be used to scale learning experiences to an individual in training based on their performance.

We propose that individuals trained with iVR will score higher in objective surgical metrics compared to those trained using a traditional technical skill instructional video. To determine this, a validated surgical outcome metric, the Objective Structured Assessment of Technical Skills (OSATS) tool, was compared as our primary measure.(Martin et al., 1997) Secondary objectives included comparison of the transfer of skill ratios (ToT), transfer effectiveness ratios (TER) with cost effectiveness, and validation of a novel iVR scoring metric, termed the Precision Score.

METHODS

The study used a single-blinded, randomized, intervention-controlled study with orthopaedic surgical residents to determine the effectiveness of iVR training in complex surgical skill acquisition.

Participants

Following Institutional Review Board approval, senior orthopaedic surgery residents (PGY4-5) who attended the 2020 Canadian Shoulder and Elbow Society (CSES) Annual Resident and Fellow Course were approached for study participation. Participants were recruited and data was collected from January 30th to February 1st, 2020. Three expert surgeons (AB, JWP, PL) who are members of the CSES were recruited to act as blinded evaluators during the study. The study was conducted at the University of Ottawa Skills and Simulation Center (uOSSC).

Twenty participants were recruited. Two individuals were excluded for not meeting inclusion criteria. Nine residents were randomized to iVR or control group training modalities, respectively. Figure 1 provides the CONSORT diagram. There were no significant differences in baseline demographic characteristics including previous VR simulator use (p=0.20) or RSA experience (p=0.81). Three residents had identified prior VR training in their surgical education. This was clarified to include previous iterations of VR simulators specifically related to arthroscopic and minimally invasive surgery. A previous review of the use of VR technology in shoulder surgery delineates this difference from previous generations of simulators using computer rendered graphics.(Lohre, Athwal, Warner, & Goel, 2020) These older iterations of VR are not the same as the immersive VR apparatus used in our study and does not incorporate head mounted display (HMD) or position tracked controllers. All of the identified VR use outside of surgical education was identified as consumer-grade VR products similar to those used in this study. Table 1 shows group demographic characteristics.

Table 1. Resident Participant Demographics

	Resident (n=18)		
	iVR Trained (Experimental) (n=9)	Video Trained (Control) (n=9)	p value
Age (mean + SD)	31.1 +/- 2.8	31.0 +/- 2.7	0.93
Gender	Male = 8 Female = 1 Undisclosed = 0	Male = 6 Female = 3 Undisclosed = 0	0.17
Post-graduate training level	PGY4 = 4 PGY5 = 5	PGY4 = 5 PGY5 = 4	0.68
Hand dominance	Left = 2 Right = 7	Left = 1 Right = 8	0.59
Any corrected vision?	Yes = 5 No = 4	Yes = 1 No = 8	0.19
Subjective experience with shoulder surgical approaches[1]	3.8 +/- 1.0	3.6 +/- 0.5	0.81
Prior experience with RSA[1]	3.3 +/- 0.9	3.2 +/- 0.4	1.0
Shoulder specific surgical courses attended[2]	1.4 +/- 0.7	1.2 +/- 0.4	0.40
Number of RSA completed acting as primary surgeon[3]	1.4 +- 0.7	1.4 +/- 0.7	0.80
Prior use of any simulator in training	Yes = 6 No = 3	Yes = 4 No = 5	0.35

Prior use of any VR products in our outside of training	Yes = 5 No = 4	Yes = 2 No = 7	0.20
Prior use of VR products in surgical training	Yes = 3 No = 6	Yes = 0 No = 9	0.08
Prior use of instructional videos in training	Yes = 9 No = 0	Yes = 9 No = 0	1.0

1 = Likert-scale 1-5
2 = Likert-scale 1-3
3 = Likert-scale 1-4

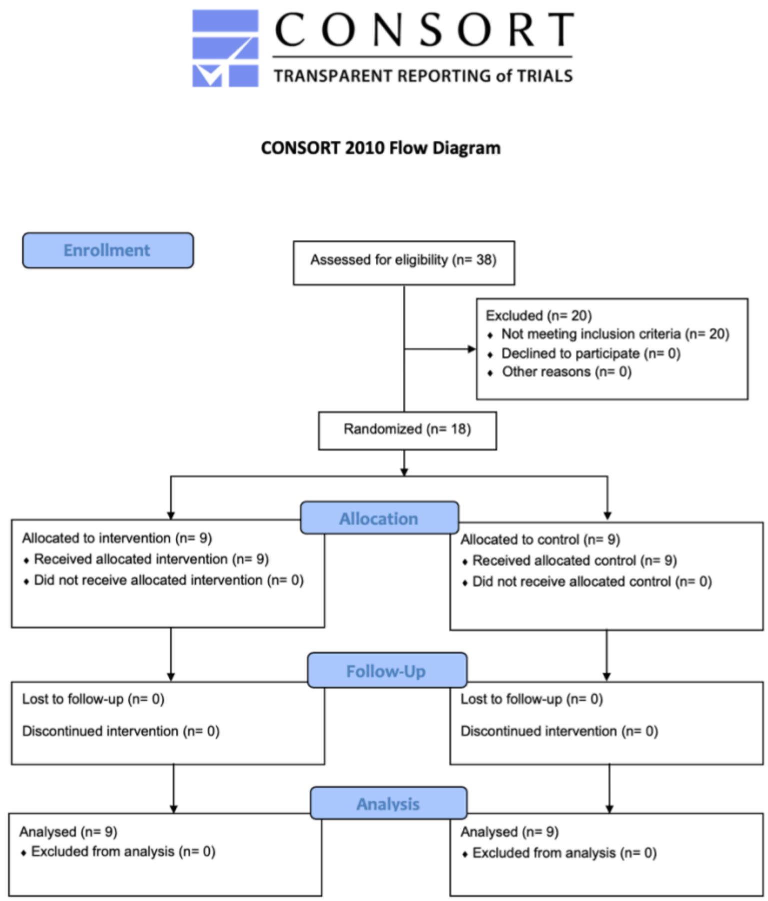

Figure 1: Recruitment and Randomization was Performed by CONSORT Recommendations

Randomization

Residents were block randomized using an internet-based, concealed computer-generated random allocation sequence. They were stratified by year of study into intervention (iVR training) or control (video training) groups. The intervention was not revealed to evaluators. Residents from both groups completed baseline Likert-scale demographic and confidence scale (CS) questionnaires.(Hecimovich, Styles, & Volet, 2014)

Training Intervention

Both groups received training on performing an RSA for rotator cuff tear arthropathy (CTA). CTA predisposes the glenoid to a superior wear pattern necessitating implants with augments.(Sirveaux, Favard, Oudet, Huquet, Walch, & Mole, 2004) The intervention (iVR) group received training on the PrecisionOS platform (version 3.0), (PrecisionOS Technology, Vancouver, BC, Canada). The iVR module provides guided learning for key steps in the procedure and provides a composite Precision Score at the end of the module. The Precision Score is calculated based on several key parameters relevant to safe and successful implantation.(Keener, Patterson, Orvets, Aleem, & Chamberlain, 2018) This includes completion time, number of repetitions, errors during key steps, and final implant orientation in three dimensional space relative to the glenoid anatomy. Essentially, the higher the Precision Score, the better the virtual surgical performance.

Control group residents received training using a video of Reverse Shoulder Arthroplasty with Augmented Baseplate.(Sperling & Zimmer Biomet, n.d.) Residents were provided with an iPad (Cupertino, USA) and headphones and were instructed to re-watch as deemed individually necessary for appropriate learning. Both groups were not limited for repetition or learning duration. A third-party research member was present during each group training activity to mitigate bias of information.

Both groups completed a written knowledge test and a repeat CS questionnaire following the training scenarios.

Secondary outcome variables were collected from the training scenarios. Training intervention variables collected included:

- Completion time of video or iVR module measured numerically in seconds
- Number of repetitions of video or iVR module measured numerically, ranging from 1-3
- Precision Score for iVR group measured numerically as a percentage, ranging from 0-100%
- Virtual reality implantation parameters of the glenoid baseplate component
    - Version – measured numerically in degrees from posterior, to neutral (0-value) to anterior
    - Inclination – measured numerically in degrees from inferior, to neutral (0-value), to superior
    - Rotation – measured numerically as a clock-face based on location of the offset, from 12 oclock (neutral, 0-value)
    - Offset – location of the centre of the glenoid baseplate from the ideal centre location measured numerically in millimeters (neutral, ideal location is 0-value)

Cadaveric Dissection and Outcomes

Fresh-frozen cadaveric specimens (scapula to hand) were prepared with a standard shoulder deltopectoral surgical approach and given a superior glenoid wear (Favard E2) pattern seen in Figure 2.(Sirveaux, Favard, Oudet, Huquet, Walch, & Molé, 2004) The superior defect was created using a standardized custom metal guide. Cadaveric specimens were mounted on clamps. A third-party assistant as well as surgical device representatives provided by Zimmer Biomet were present to act as technicians and assist with surgical equipment. All assists were instructed to limit interaction with study participants and not offer any advice.

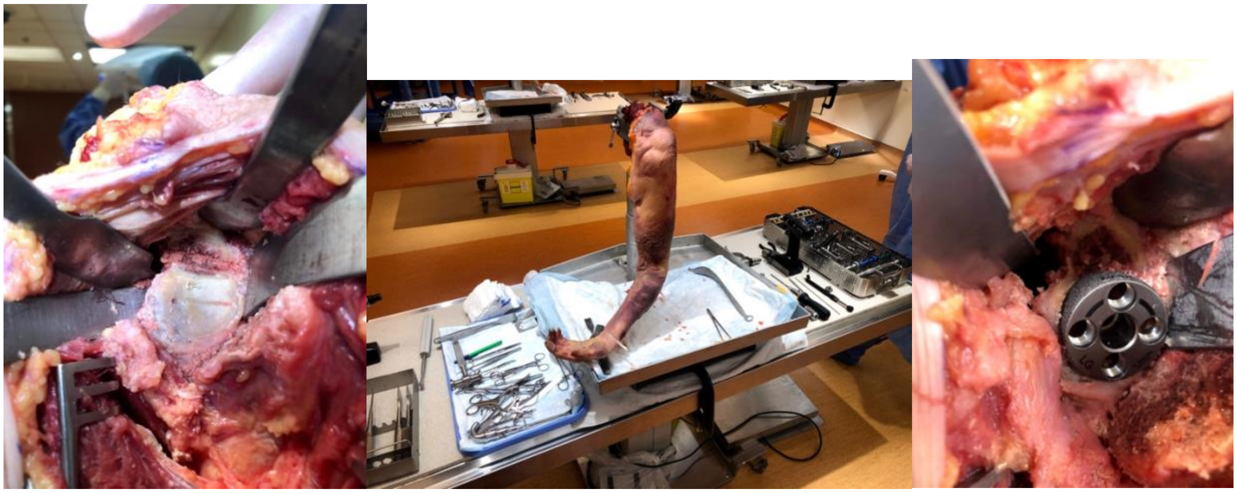

Figure 2. (a) Representative Cadaveric Specimen with Imparted Favard E2 Wear Pattern Produced by Expert Surgeons (JP, AB) Using Standardized Instrumentation (b) An Evaluation Station, with Fresh-Frozen Cadaveric Specimen and Instrumentation Sets. (c) A Final Implanted Reverse Shoulder Arthroplasty with Augment Placed Into Superior Defect

Each participant was brought to a cadaveric lab to be assessed by blinded evaluators for augmented glenoid implantation. Residents were asked verbal questions and were timed for their responses while performing the dissection. Questions included identifying the location of the cadaveric wear pattern, predisposing condition to this pathology, visually identifying the wear pattern on a series of illustrations, describing the appropriate position of guide pin scapular exit, and ideal ream depth for the wear pattern. Residents were instructed on the role of assistants. The residents then performed an RSA using a Zimmer Biomet Comprehensive Reverse Shoulder Arthroplasty with Augmented Baseplate system (Warsaw, United States). Residents were timed for task completion, starting from asking for or touching first instrument, to final glenoid baseplate implantation. Prior to evaluating the participants' performance, the blinded assessors received training on the evaluation tools including the Objective Structured Assessment of Technical Skills (OSATS) and Global Ratings Scale (GRS). These are validated metrics of performance commonly used in procedural education assessments. The OSATS score is a composite check list of sequential and key actions during the procedure. Each task is given a value of 0 or 1 based on appropriate completion. The GRS is a subjective assessment of performance in domains of respect for tissue, time and motion efficiency, instrument handling, knowledge of instrumentation, flow of procedure, knowledge of procedure, and an overall score. The GRS questions are in Likert-scale format, scored 1-5. Evaluators evaluated residents using OSATS and GRS. The quality of final products made in orthopaedics are not necessarily reflected by these evaluation metrics. Therefore, residents and evaluators were asked to provide subjective final implant parameters of version, inclination, rotation, and offset to assess product quality. These questions were asked to participants and evaluators separately to avoid bias.

Residents then completed final Likert-scale questionnaires consisting of learning activity assessment in parameters of realism, teaching capacity, and perceived longitudinal benefit, and subjective/self-assigned GRS scores. Table 2 provides outcome measures and their descriptions.

Table 2. Primary and Secondary Outcome Measures Used in the Study with Descriptors

Outcome	Measure	Description	Variable Type	Range	Calculation	Interpretation
Primary	OSATS	Composite check list of key and sequential steps in a	Ordinal	Each step 0 or 1. A "1" indicates completion of the step. Composite	Addition of each step to determine sum.	A higher OSATS score indicates more appropriately completed

		surgical procedure		step score of 20 for this experiment		key procedural steps, and therefore more understanding and technical ability to perform the procedure as a whole
Secondary	GRS	Performance assessment in domains of: (1) respect for tissue, (2) time and motion, (3) instrument handling, (4) knowledge of instruments, (5) flow of procedure, (6) knowledge of specific procedure, (7) overall performance, (8) quality of final product	Ordinal	Each domain consists of a Likert-scale of 1-5. A higher score indicates greater alignment with better performance in that domain. Composite scores can be produced of each domain together.	Addition of each step to determine sum.	Each domain can be assessed individually, or collectively. A higher score in each domain or collectively, respectively, indicates improved performance. When evaluated using this by an expert, higher scales indicate higher performance more in line with expert expectations.
Secondary	Transfer of training (ToT)	A comparison of skill achievement by those trained using the experimental simulation versus a control	Numerical – expressed as percent	A negative value indicates simulator training detracts from skill acquisition. A value of 0% indicates no effect of the simulator. A value of 100% indicates that an individual could reach proficiency	$ToT = \left(\frac{T_{video(cadaver)} - T_{iVR(cadaver)}}{T_{video(cadaver)}}\right) \times 100\%$	The percentage value indicates how much skill is gained using simulation training compared to the control group. A higher value indicates less dependency on real-world training and a reduction in learning curve of procedures.

				only using the simulator, not requiring any additional training.		
Secondary	Transfer Effectiveness Ratio (TER)	A comparison of skill achievement incorporating task completion time	Numerical	Values from 0 to beyond 1.0. Values of 1.0 or greater indicate training is more effective in simulator than real life	$TER = \left(\frac{T_{video(cadaver)} - T_{iVR(cadaver)}}{T_{iVR(simulated)}}\right)$	Indicates how much real-world training time is saved by using a simulator.
Secondary	Cost-Effectiveness Ratio (CER)	A cost saving comparison between simulator training and traditional training	Numerical	Values from 0 upwards. Higher values indicate greater cost savings	$CER = \frac{TER}{\left(\frac{Cost(iVR)}{Cost(control)}\right)}$	Indicates how much money is saved using a simulator compared to a control by incorporating training time and TER
Secondary	Confidence Scale (CS)	A developed psychometric scale adapted from (Grundy, S. E. (1993). The Confidence Scale. Nurse Educator, 18(1), 6–9.)	Ordinal	Likert-scale of 1-5, consisting of 6-questions. The higher the value, the more confident the performance	n/a	Higher cumulative scores for each Likert-type question corresponds to increasing confidence in performance
Secondary	Perception survey	Questions pertaining to: Enjoyment Realism Teaching capacity	Ordinal	Likert-scale questions from 1-5. Higher values indicate greater performance in each domain		Enjoyment: higher scores indicate greater enjoyment Realism: higher scores indicate greater realism and thus, face validity Teaching capacity: higher scores indicate greater ability

						to teach and thus, content validity
Secondary	Knowledge scores: Verbal Written	Questions pertaining to shoulder replacement principles and specific pathology identification	Numerical	Cumulative scoring from 0 to 16 for written, and 0 to 6 for verbal	-	Higher scores indicate improved knowledge and recall.

Precision Score

The Precision Score is a virtual metric of performance similar to real-world metrics such as OSATS or GRS. The Precision Score however accounts for final, virtual product orientation parameters based on best available evidence. The Precision Score is automatically generated by the software based on user performance in completing or not completing key procedural virtual steps. The Precision Score for the RSA is provided by time of each key step, appropriate completion of each step, and the 3-dimensional performance quality of each step. The final score is a composite of these performance areas including quality of final product. The Precision Score is provided to the user and was collected for comparison to real-world performance.

Outcomes and Statistical Analysis

The primary outcome consisted of OSATS score to determine any superiority using iVR training compared to control for learning decision making and technical skills in complex RSA. Secondary outcomes include GRS score, ToT, TER, and CER scores of iVR training relative to control as well as validation of the Precision Score, a novel virtual reality based ratings scale.
To achieve 80% statistical power (beta=0.2) for the primary outcome measure (OSATS), a 2-sided test at alpha=0.05 showed a minimum of six subjects were required for each group based on a conservative estimate of 25% difference in combined outcome measures of iVR training.(Lohre, Bois, et al., 2020) Data was tested for normality prior to statistical analysis and analyzed by the “intention-to-treat principle.” Student’s t-test was performed for 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. Pearson product correlation was used to determine similarity and correlation between ratings scales. Cronbach’s alpha was utilized to determine reliability of Likert scales. Results were considered significant at a p<0.05. Data was handled as a complete-case analysis.

RESULTS

Table 3. Key, Statistically Significant Study Results for Primary and Secondary Objectives

	iVR trained (experimental; n=9)	Video trained (control; n=9)	Confidence interval and p-value	Cumulative Value	Interpretation
Primary Objective OSATS	M = 15.9, SD = 2.5; 95% CI [14.3,17.5]	M = 9.4, SD = 3.2; 95% CI [7.4,11.5]	95% CI [3.3,9.7]; p=0.0007	-	Cohen’s d effect size = 2.26 (very strong). The majority of control group performance fell below the mean of the iVR trained group

Secondary Objective Verbal Knowledge Scores	M = 4.1, SD = 1.0; 95% CI [3.4,4.8]	M = 2.2, SD = 1.7;95% CI [1.1,3.3]	95% CI [3.3,9.7]; p=0.0007	-	Cohen's d effect size = 1.9 (very strong). The majority of control group performance fell below the mean of the iVR trained group
Secondary Objective Training Time (s)	M = 10.4, SD = 5.0 (95% CI [7.10,13.7]	M = 16.1, SD = 2.6 (95% CI [14.4,17.8]	p=0.008; 95% CI [1.64,10.3]	-	The iVR group trained 155% faster than the control group
Secondary Objective Transfer of Training (ToT)	-	-	-	32.6-59.4%	Calculated from cadaveric completion time and OSATS scores (accounting for value range). Based on early learning curve studies in shoulder replacement, training using iVR could account for reductions in real-world training time of up to 51 cases.
Secondary Objective Transfer Effectiveness Ratio (TER)	-	-	-	0.79	One-hour of iVR training reduces real-world training time by 48 minutes
Secondary Objective Cost-Effectiveness Ratio (CER)	-	-	-	37	Compared to a traditional course model including flights, accommodation, fees, iVR training amounts to $10.95/day and is approximately 37x more cost effective based on performance of learning
Secondary Objective	M = 0.75, SD = 0.13 (range 0.52	-	-	Correlation to:	The virtual performance

Precision Score	to 0.92), max 1.0			OSATS r(8) = 0.74, p = 0.035 (strong) Implant orientation parameters r(12) = 0.73, p = 0.006 (strong)	metric had strong correlations to real-world performance, indicating potential future use as a metric of tracking performance and learning

Resident Experience

Residents trained in iVR had greater enjoyment of learning (95% CI [0.20,2.6]; p=0.01), felt it was easy to use (95% CI [0.003,1.3]; p=0.02), and provided a greater capacity for teaching compared to the video group (95% CI [0.3,2.2]; p=0.01). This included domains of anatomy teaching (95% CI [0.5,2.2]; p=0.002), general (95% CI [0.1,2.1]; p=0.009) and implant specific (95% CI [0.8,1.6]; p=0.001) surgical steps.

Training Intervention

Residents trained in iVR completed their training session 155% faster than control (p=0.008; 95% CI [1.64,10.3]. Mean iVR training time was 10.4, SD = 5.0 (95% CI [7.10,13.7]) minutes versus M = 16.1, SD = 2.6 (95% CI [14.4,17.8]) minutes for video training. Residents completed 2-3 module repetitions during iVR training. There was a significant difference in module completion time between trials (95% CI [3.41,8.90]; p=0.0005), with an average of M = 6.68, SD = 3.12 minutes (range 1.6 to 10.4 minutes; 95% CI[4.52,8.85]) reduction between first and second trial. The CS showed no significant difference between groups for baseline (p=0.50) or post-training confidence (p=0.78) with an internal consistency of 0.88. However, a greater number of iVR trained residents (n=4) had positive increases in their CS scores compared to video (n=3) and one video trained resident actually had lower CS scores (p=0.10).

Cadaveric Dissection

Immersive VR trained residents demonstrated significantly higher OSATS scores (M = 15.9, SD = 2.5; 95% CI [14.3,17.5]) than the video group (M = 9.4, SD = 3.2; 95% CI [7.4,11.5])(95% CI [3.3,9.7]; p=0.0007). Immersive VR trained residents completed the cadaveric procedure faster (M = 17.1, SD = 5.7 minutes; 95% CI[13.4,20.8]) than the video group (M = 25.3, SD = 32.5 minutes; 95% CI[4.1,46.5]) (95% CI [-39.4,21.0]; p=0.13). Global Rating Scale total and individual domain scores were not significantly different between groups (p=0.92). There were no major implant specific complications in both groups. Immersive VR training time moderately predicted real-world implantation time (r(9) = 0.42, p = 0.25) compared to control training time (r(9) = 0.04, p = 0.92).

Residents trained in iVR showed significantly higher verbal questioning scores (M = 4.1, SD = 1.0; 95% CI [3.4,4.8]) than those video trained (M 2.2, SD = 1.7;95% CI [1.1,3.3])(95% CI [0.10,3.3];p=0.03). Resident groups had no significant differences in written knowledge scores (p=0.50).

Residents receiving iVR training were provided with a Precision Score, computed from composite data of implantation time and orientation parameters in the virtual patient environment. Mean Precision Score was 0.75, SD = 0.13 (range 0.52 to 0.92), max 1.0. There was no significant change between Precision Scores between module completions (p=0.81). The Precision Score showed a strong correlation (r(8) = 0.74, p = 0.035) to OSATS scores. The Precision Score had "good" internal consistency of 0.82. Precision Score was also correlated to GRS scores (r(8) = 0.32, p = 0.44) and completion time (r(8) = 0.43, p = 0.29). The Precision Score was seen to correlate strongly with implant orientation parameters provided by expert surgeons (r(12) = 0.73, p = 0.006) for the final construct. Figure 3 shows a representative Precision Score and example of the virtual training environment.

Figure 3. (a) Immersive Virtual Reality Operating Room for Case-Based Learning and Learning Reverse Shoulder Arthroplasty (b) Immediate Metric Feedback Used to Calculate the User's Precision Score

Training Validity

The ToT, TER, and CER ratios were calculated for the iVR training relative to control training. The range of ToT ratios was 32.6-59.4% based on cadaveric dissection time and OSATS scores, respectively. The TER ratio was calculated as 0.79 when comparing performance of iVR training to control. Figure 4 demonstrates the effects of ToT on early learning curves. The cost incurred for iVR training was considered a function of daily use on a $4000.00 per year cost, equaling $10.96/day, or in this case, per training session. The cost of traditional training was considered the cost incurred to an individual for course registration and travel and was approximated as $1000.00. Based on these estimates, the CER was 37.0.

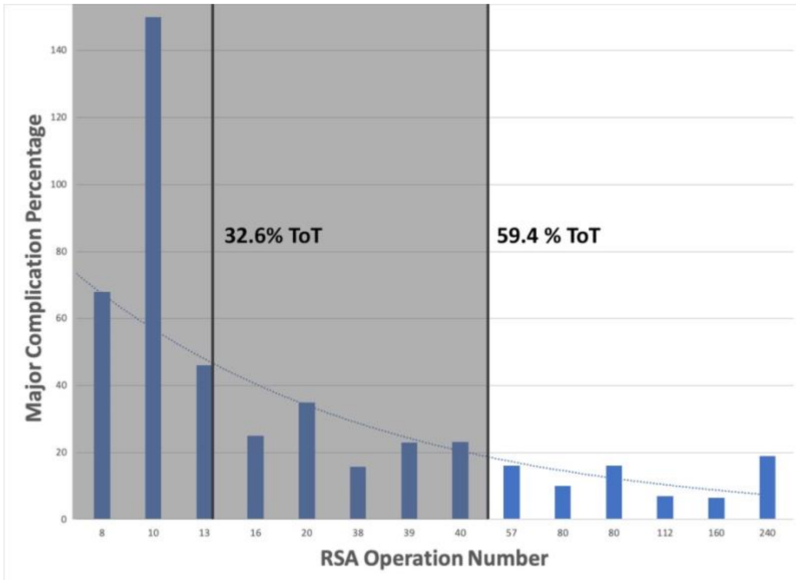

Figure 4. Fourteen representative studies of early learning curves and composite major complication rates (expressed as percentages) for performing reverse shoulder arthroplasty. A fitted curve (r^2 = 0.80) allows for determination of early learning curve reduction based on transfer of training ratios. Our ToT of 32.5-% - 59.4% illustrates iVR training reduces early learning curves by 13 to 51 cases, represented by grey rectangle and solid vertical lines.

DISCUSSION

By means of a randomized, intervention-controlled trial, iVR was demonstrated to be superior to technical video training for acquisition of procedural knowledge, as well as pathology recognition and decision making. The iVR trained group had superior OSATS scores as well as higher verbal questioning scores following a single training session.

Human performance evaluation requires consistent, validated metrics. In surgical education research, the performance metric used is dependent on the type of surgery performed. Considering our procedure is an open approach with implantable materials, the use of OSATS and GRS are the most widely used and accepted methods of evaluation (Kohls-Gatzoulis, Regehr, & Hutchison, 2004). Human performance however is a composite of skill and knowledge. As such, we devised our study to incorporate both written questions and interspersed verbal questions occurring before, during, and after the cadaveric dissection. These verbal questions were timed. Human performance is affected by stressful situations and incorporating timed events with verbal questioning aside from the task at hand adds to this stress (Sanders, 1983). Using transfer of training (ToT) and effectiveness (TER) ratios, one can determine whether the simulator in fact accomplishes the training task for which it was developed (Roscoe, 1971). This method of transfer validity has been employed in medical education, surgical education, and in other industries including aeronautics and military application. A number of control participants in our study rapidly completed the glenoid implantation task albeit incorrectly due to missing many steps. This accurately describes the importance of using these ratios rather than simple metrics or time to completion alone.

Humans learn more effectively through multisensory experiences often including visual, auditory and kinesthetic.(Shams & Seitz, 2008) Immersive VR uniquely provides this learning environment, and this may be reflected in both the improved knowledge and skill. Educational literature provided by Edgar Dale and Anders Ericsson shows that combining informational learning with technical tasks improves performance.(Lee & Reeves, 2007),(K. A. Ericsson, 2004) The term deliberate practice emphasizes focused, cognitive appreciation of the task performed by the trainee. This has been shown to help achieve “expert” status in multiple fields of human performance and is reliant on three features: (1) immersive experience, (2) personalized metric involvement, (3) permission of failure with reflective repetition.(A. Ericsson & Harwell, 2019) The iVR training environment uniquely satisfies these principles. Additionally, iVR allows for learning outside of potential patient harm and in remote settings. This can allow for learning outside of traditional schemes of courses and clinical rotations – extending reach and adding significant value to skill acquisition and retention strategies. Particularly for military medical personnel, this could mean extended training during deployment and an ability to practice infrequently used skills outside of their particular subspecialty area. Skill decay during deployment is of concern.(Arthur & Day, 2019) Skill decay and reacquisition are directly proportional in a study examining military surgeons deployed up to 6 months in duration.(Perez et al., 2013) Consistent, virtual psychomotor training in iVR may reduce skill decay during deployment periods, or alternatively assist in reacquisition.

Our study demonstrated a ToT of 32.6-59.4% based on cadaveric dissection time and OSATS scores, respectively. Transfer of training informs potential reduction in early learning curves by training with a simulator. Our study incorporated a complex surgical task in an elective orthopaedic procedure. There is inconsistency in defining an early learning curve for reverse shoulder arthroplasty with some studies showing high complication rates into the range of 150 cases.(Gilles Walch, Bacle, Lädermann, Nové-Josserand, & Smithers, 2012) Based on available evidence for early learning curves in RSA from multiple, experienced surgical groups, the ToT achieved using the iVR simulator would account for performing 13-51 RSA.(Boulahia, Edwards, Walch, & Baratta, 2002; Choi, Bae, Kwon, & Kang, 2019; De Wilde, Plasschaert, Audenaert, & Verdonk, 2005; Gallo et al., 2011; Groh & Groh, 2014; Keener et al., 2018; Rittmeister & Kerschbaumer, 2001; Sirveaux, Favard, Oudet, Huguet, & Lautman, 2001; Sirveaux, Favard, Oudet, Huquet, Walch, & Mole, 2004; Valenti, Boutens, & Nerot, 2001; Gilles Walch et al., 2012; Wierks, Skolasky, Hun Ji, & McFarland, 2009) The large range is accounted for by the published range of case learning, as well as calculation methods based on OSATS scores and procedural time, and likely favors the more conservative value. Though this has been demonstrated using an elective surgical procedure, we feel that the learning modality and principles can be translated to other surgical procedures.

A unit perspective of combat casualties during Operation Iraqi Freedom by the U.S. Army shows explosive injuries accounting for 81% of casualties.(Belmont et al., 2010) Musculoskeletal injuries account for 49.4% of combat casualties.(Belmont et al., 2010) Died of wounds (DOW) or MEDEVAC rates in this study were 23.7 per 1000 soldier combat-years.(Belmont et al., 2010) This spectrum of injury and improved survival and transfer rates compared to previous combat engagements signifies the importance of competent surgical stabilization in the field. Consistent training methods incorporating feedback and principles of deliberate practice such as iVR can extend training reach. Current methods of military training at level I trauma centers is not consistent in exposure to battlefield injury patterns.(Breeze et al., n.d.) VR has shown patterns of sympathetic, or stress response to immersive environments.(Varela-Aldas, Placios-Navarro, Garcia-Magarino, & Fuentes, 2019) Virtual training in an immersive environment can replicate these injuries including the stressful patterns of care and aside from actual battlefield exposure, will likely be the closest training scenario to real-world operations.(Fan & Wen, 2019)

TER accounts for time spent in the simulator compared to training time in the control environment, relative to real-world procedural time to reach task completion.(Roscoe, 1971) Transfer effectiveness ratio additionally provides information on potentially reduced training times by using a simulator. A TER of 1.0 indicates simulator-based training is equivalent to real-world training. Interpreting our TER of 0.79 shows that 1 hour of iVR training saves 48 minutes of real-world training time. The reported average flight simulator TER is 0.33, including those that are used for licensing.(Yoon, Park, Lee, & Kim, 2019) One of the most prominently studied VR simulators is the MIST-VR laparoscopic trainer, which has previously shown a TER of 0.42.(Gallagher, Seymour, Mcglade, & Satava, 2013) Our study employed an experimental-versus-control-group method of ToT and TER determination, considered the most appropriate for validity determination.(Caro, 1977),(Korteling, Oprins, & Kallen, 2009) Immersive VR simulators have already been employed in the field to train infantry.(Fan & Wen, 2019) These units consist of simulated, team environments to practice communication and battlefield tactics. The hardware involves inertial sensor suits and are costly. Extending the use of iVR technology to medical personnel would provide similar, consistent training outside of dangerous environments.

From a cost perspective, one minute of OR time costs $37USD.(Childers & Maggard-Gibbons, 2018) Considering the reduction in on the job learning time provided by the iVR construct, this has the potential to greatly influence procedural training time and costs. Safely reducing learning curves in a virtual environment reduces complication costs. Immersive VR training allowed residents to more correctly identify surgical pathology and provide an appropriate surgical construct based on available instruments in the surgical tray during the study. This represents high level learning beyond simple task repetition and has never been previously shown in immersive VR simulators. We have shown a CER based on improved training time provided by the TER, relative to video training, and the cost of a single training course as a surrogate of traditional training methods. Cost considerations were performed per recommendations by Fletcher and Sottilare.(Fletcher & Sottilare, 2014) With these assumptions, the iVR training is at minimum 37 times more cost effective than our control. If one considers attending multiple courses, or incorporating the per minute OR cost, the cost effectiveness significantly increases. Bioskills courses including cadaveric specimens are costly. Total course fees including travel extend into the $6000 USD range for surgeons.(Kelliher, 2012) Individual cadaveric specimens additionally are costly. Our study cost related to cadaveric specimens only was over 8000 USD including specimens, equipment trays, clamps, rental costs and transportation. Additional incurred costs are that of support staff, venue, and disposal costs. Following the COVID-19 pandemic, training costs are expected to increase due to staff furloughing and disruptions in supply chains.(Menendez, Jawa, Haas, Warner, & The Codman Shoulder Society, 2020) Figure 5 shows costs to a single surgeon for different continuing medical education strategies. It is not unreasonable to expect an additional 1000 USD to 2000 USD added to the cost of bioskills courses, meetings, or site visits. In comparison, the daily cost of the iVR simulator at 10.96/day USD is significantly more cost effective given demonstrable skill and knowledge acquisition. It is also not expected to increase due to the current pandemic financial disruptions. A formal return on investment (ROI) study is currently being performed in collaboration with Jack Philips and the ROI Institute.

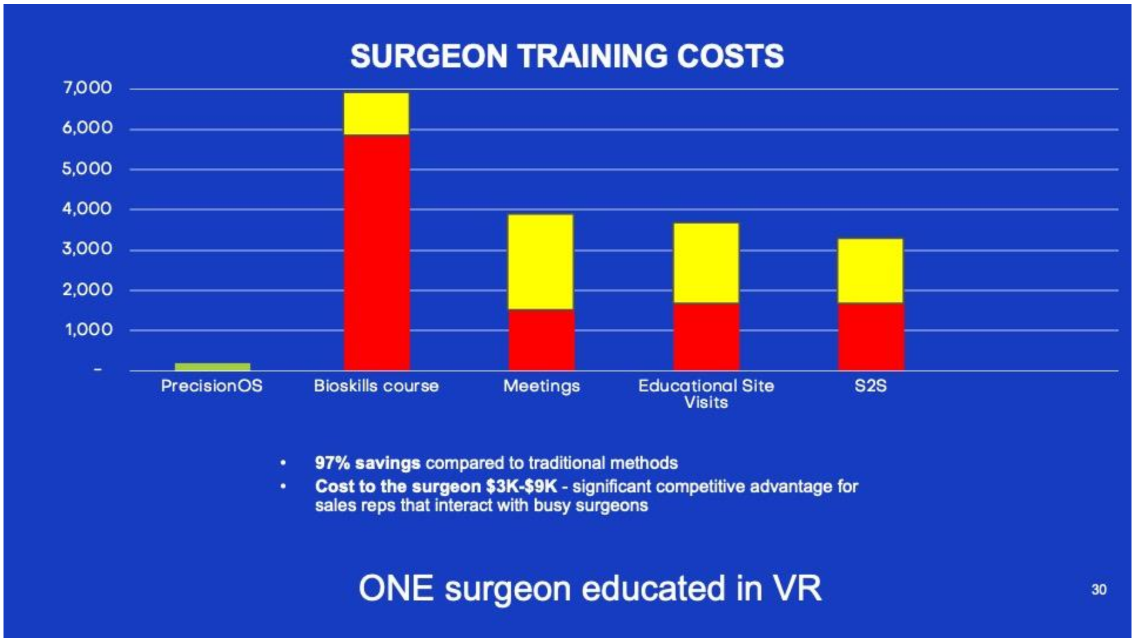

Figure 5. Training Costs of a Single Surgeon Using the Spectrum of Continuing Medical Education Modalities. Red Bars Represent Current Costs and Yellow Represent Expected Increased Costs Following the COVID-19 Pandemic

Validated scoring metrics are available for open and minimally invasive procedures. These include the OSATS, GRS and ASSET. These have previously been validated for skill acquisition and demonstration in real world scenarios.(Jacobsen, Anderson, Hansen, & Konge, 2015) To date, there is no rating scale for iVR given its novelty. The Precision Score was developed to determine training module outcome in the virtual world. One benefit of iVR is the ability to provide immediate metrics and feedback to users. The Precision Score is an adaptable score that incorporates time to task completion with evidence-based parameters of achievement including implant positioning. This is unique over most real-world metrics of performance as they typically do not assess quality of final product.(Anderson et al., 2016) Furthermore, the score is produced by software using the iVR system and is not reliant on individual assessments. Our study is the first of its kind to validate the use of a VR ratings scale. The Precision Score measured strong correlation coefficients and internal consistency compared to OSATS performance. Importantly in procedural applications, the Precision Score performance also strongly correlated to real-world improvements in implantation parameters. This adaptable score must be further used and validated to become the standard of outcomes measurement with the increasing use of iVR technology.

Study limitations include small recruitment numbers based on convenience sampling and availability. The confidence scale used may be too granular for single training sessions given task complexity. Our ToT, TER and CER does not inform longitudinal benefit of the iVR training environment. An incremental transfer effectiveness ratio (ITER) calculated from multiple training sessions would inform the loss of real-world training time with simulator repetition. It was not possible to calculate this in the current study given our study criteria but should be examined in the future. We used cadaveric specimens instead of a real operative scenario. However, using cadavers has been demonstrated as an appropriate substitute.(Bjerrum et al., 2017) Consistency of scoring between evaluators could be improved with intervention recording and reassessment by multiple evaluators. We have additionally compared iVR training to technical video training alone, and not mixed media methods which are commonly employed by learners.

CONCLUSION

Learning complex procedural skills in iVR was found to be superior to technical video training. The iVR training platform provided improved knowledge and pathology recognition to senior surgical residents in a single session. Validated, objective measures of training effectiveness showed reduction in both theoretic early learning curves and real-world training time with iVR use. The TER value for iVR was more significant than prominent surgical simulators previously examined. The newly developed Precision Score, an iVR scoring metric, correlated with both technical skill in the real-world task and final product quality, thus simultaneously providing and validating a novel assessment tool for the virtual simulation environment. Further research into the effects of longitudinal iVR learning must be undertaken.

ACKNOWLEDGEMENTS

We would like to acknowledge the Canadian Orthopaedic Association (COA) and the Canadian Shoulder and Elbow Society (CSES) for providing funding and operational support during the study.

REFERENCES

Anderson, D. D., Long, S., Thomas, G. W., Putnam, M. D., Bechtold, J. E., & Karam, M. D. (2016). Objective Structured Assessments of Technical Skills (OSATS) Does Not Assess the Quality of the Surgical Result Effectively. *Clinical Orthopaedics and Related Research*, *474*(4), 874–881. https://doi.org/10.1007/s11999-015-4603-4

Arthur, W. J., & Day, E. A. (2019). Skill Decay: The Science and Practice of Mitigating Skill Loss and Enhancing Retention. In P. Ward, J. M. Schraagen, J. Gore, & E. M. Roth (Eds.), *The Oxford Handbook of Expertise*. Oxford University Press.

Bartlett, J. D., Lawrence, J. E., Stewart, M. E., Nakano, N., & Khanduja, V. (2018). Does virtual reality simulation have a role in training trauma and orthopaedic surgeons? *Bone and Joint Journal*, *100B*(5), 559–565. https://doi.org/10.1302/0301-620X.100B5.BJJ-2017-1439

Belmont, P. J., Goodman, G., Zacchilli, M., Posner, M., Evans, C., & Owens, B. (2010). Incidence and epidemiology of combat injuries sustained during “the surge” portion of operation Iraqi Freedom by a U.S. Army brigade combat team. *J Trauma*, *68*(1), 204–210.

Bjerrum, F., Konge, Ã. L., Subhi, Ã. Y., & Thomsen, A. S. S. (2017). Gathering Validity Evidence for Surgical Simulation A Systematic Review. *Annals of Surgery*, *XX*(Xx), 1–6. https://doi.org/10.1097/SLA.0000000000002652

Boulahia, A., Edwards, T., Walch, G., & Baratta, R. (2002). Early results of a reverse design prosthesis in the treatment of arthritis of the shoulder in elderly patients with a large rotator cuff tear. *Orthopedics*, *25*, 129–133.

Breeze, J., Combes, J., DuBose, J., & Powers, D. (n.d.). How are we currently training and maintaining clinical readiness of US and UK military surgeons responsible for managing head, face, and neck wounds on deployment? *2018*, *164*(3), 183–185.

Caro, F. G. (1977). *Readings in Evaluation Research* (2ed ed.). New York: Russell Sage Foundation.

Childers, P. C., & Maggard-Gibbons, M. (2018). Understanding Costs of Care in the Operating Room. *JAMA*

Surgery, *90095*. https://doi.org/10.1001/jamasurg.2017.6233
Choi, S., Bae, J., Kwon, Y. S., & Kang, H. (2019). Clinical outcomes and complications of cementless reverse total shoulder arthroplasty during the early learning curve. *Journal of Orthopaedic Surgery and Research*, *14*(53).
De Wilde, L., Plasschaert, F., Audenaert, E., & Verdonk, R. (2005). Functional recovery after a reverse prosthesis for reconstruction of the proximal humerus in tumor surgery. *Clin Orthop Relat Res*, *430*, 156–162.
Ericsson, A., & Harwell, K. (2019). Deliberate Practice and Proposed Limits on the Effects of Practice on the Acquisition of Expert Performance: Why the Original Definition Matters and Recommendations for Future Research. *Front Psychol*, *10*, 2396.
Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine : Journal of the Association of American Medical Colleges*, *79*(10 Suppl), S70-81.
Fan, Y.-C., & Wen, C.-Y. (2019). A Virtual Reality Soldier Simulator with Body Area Networks for Team Training. *Sensors (Basel)*, *19*(3), 451.
Fletcher, J., & Sottilare, R. A. (2014). Cost Analysis for Training and Education Systems. In R. A. Sottilare, A. Graesser, X. Hu, & B. Goldberg (Eds.), *Design Recommendations for Intelligent Tutoring Systems - Volume 2: Instructional Management* (pp. 3–18). US Army Research Laboratory.
Gallagher, A. G., Seymour, N. E., Mcglade, K., & Satava, R. M. (2013). Prospective , Randomized Assessment of Transfer of Training (ToT) and Transfer Effectiveness Ratio (TER) of Virtual Reality Simulation Training for Laparoscopic Skill Acquisition, *257*(6), 1025–1031. https://doi.org/10.1097/SLA.0b013e318284f658
Gallo, R., Gamradt, S., Mattern, C., Cordasco, F., Craig, E., Dines, D., … Sports Medicine and Shoulder Service at the Hospital for Special Surgery, New York, N. (2011). Instability after reverse total shoulder replacement. *J Shoulder Elbow Surg*, *20*(4), 584–590.
Groh, G. I., & Groh, G. M. (2014). Complications rates, reoperation rates, and the learning curve in reverse shoulder arthroplasty. *Journal of Shoulder and Elbow Surgery*, *23*(3), 388–394. https://doi.org/10.1016/j.jse.2013.06.002
Hecimovich, M. D., Styles, I., & Volet, S. E. (2014). Development and psychometric evaluation of scales to measure professional confidence in manual medicine: a Rasch measurement approach. *BMC Research Notes*, *7*(338).
Jacobsen, M. E., Anderson, M. J., Hansen, C. O., & Konge, L. (2015). Testing Basic Competency in Knee Arthroscopy Using a Virtual Reality Simulator: Exploring Validity and Reliability. *J Bone Joint Surg (Am)*, *97*, 775–781.
Keener, J., Patterson, B., Orvets, N., Aleem, A., & Chamberlain, A. (2018). Optimizing reverse shoulder arthroplasty component position in the setting of advanced arthritis with posterior glenoid erosion: a computer-enhanced range of motion analysis. *J Shoulder Elbow Surg*, *27*(2), 339–349.
Kelliher, J. (2012). Bioskills Solutions offers continuing education on wheels for area medical professionals.
Kohls-Gatzoulis, J. A., Regehr, G., & Hutchison, C. (2004). Teaching cognitive skills improves learning in surgical skills courses: A blinded, prospective, randomized study. *Canadian Journal of Surgery*, *47*(4), 277–283.
Korteling, J. E. H., Oprins, E. A. P. B. E., & Kallen, V. L. (2009). Measurement of Effectiveness for Training Simulations. *NATO RTO*, *SAS-095*(2005), 1–12.
Lee, S. J., & Reeves, T. C. (2007). Edgar Dale: A significant contributor to the field of educational technology. *Educational Technology*, *47*(6), 56.
Logishetty, K., Gofton, W., Rudran, B., Beaule, P., & Cobb, J. (2020). Fully Immersive Virtual Reality for Total Hip Arthroplasty. *J Bone Joint Surg (Am)*, *102*(6), e27.
Lohre, R., Athwal, G. S., Warner, J. P., & Goel, D. P. (2020). The Evolution of Virtual Reality in Shoulder and Elbow Surgery. *JSES Open Access*.
Lohre, R., Bois, A., Athwal, G. S., Goel, D. P., & (CSES), T. C. S. and E. S. (2020). Improved Complex Skill Acquisition by Immersive Virtual Reality Training: A Randomized Controlled Trial. *J Bone Joint Surg (Am)*, *102*(6), e26.
Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *The British Journal of Surgery*, *84*(2), 273–278.
Menendez, M., Jawa, A., Haas, D., Warner, J. J., & The Codman Shoulder Society. (2020). Orthopedic surgery post COVID-19: an opportunity for innovation and transformation. *Journal of Shoulder and Elbow Surgery*, 1–18.
Perez, R. S., Skinner, A., Weyhrauch, P., Niehaus, J., Lathan, C., Schwaitzberg, S. D., & Cao, C. G. L. (2013). Prevention of Surgical Skill Decay, *178*, 76–87. https://doi.org/10.7205/MILMED-D-13-00216
Rittmeister, M., & Kerschbaumer, F. (2001). Grammont reverse total shoulder arthroplasty in patients with

rheumatoid arthritis and nonreconstructable rotator cuff lesions. *J Shoulder Elbow Surg*, *10*, 17–22.

Roscoe, S. N. (1971). Incremental Transfer Effectiveness. *Technical Report ARL-70-5/ AFOSR-70-1*.

Sanders, A. (1983). Towards a model of stress and human performance. *Acta Psychologica*, *53*(1), 61–97.

Shams, L., & Seitz, A. (2008). Benefits of multisensory learning. *Trends Cogn Sci*, *12*(11), 411–417.

Sirveaux, F., Favard, L., Oudet, D., Huguet, D., & Lautman, S. (2001). Grammont inverted total shoulder arthroplasty in the treatment of glenohumeral osteoarthritis with massive nonrepairable cuff rupture. In G Walch, P. Boileau, & D. Molé (Eds.), *Shoulder Prosthesis: Two to Ten Year Follow-up* (1ed ed., pp. 247–252). Montpellier, France: Sauramps Medical.

Sirveaux, F., Favard, L., Oudet, D., Huquet, D., Walch, G., & Mole, D. (2004). Grammont inverted total shoulder arthroplasty in the treatment of glenohumeral osteoarthritis with massive rupture of the cuff: results of a multicentre study of 80 shoulders. *J Bone Joint Surg Br*, *86*, 388–395.

Sirveaux, F., Favard, L., Oudet, D., Huquet, D., Walch, G., & Molé, D. (2004). Grammont inverted total shoulder arthroplasty in the treatment of glenohumeral osteoarthritis with massive rupture of the cuff. *J Bone Joint Surg (Am)*, *86*(3), 388–395.

Sperling, J., & Zimmer Biomet. (n.d.). Reverse Shoulder with Mini Humeral Tray and Augmented Baseplate.

Valenti, P., Boutens, D., & Nerot, C. (2001). Delta 3 reversed prosthesis for osteoarthritis with massive rotator cuff tear: long-term results (>5 years). In G Walch, P. Boileau, & D. Molé (Eds.), *Shoulder Prosthesis: Two to Ten Year Follow-up* (1ed ed., pp. 253–259). Montpellier, France: Sauramps Medical.

Varela-Aldas, J., Placios-Navarro, G., Garcia-Magarino, I., & Fuentes, E. M. (2019). Effects of Immersive Virtual Reality on the Heart Rate of Athlete's Warm-Up. In L. De Paolis & P. Bourdot (Eds.), *Lecture Notes in Computer Science, vol 11613* (eds). Spring, Cham.

Walch, Gilles, Bacle, G., Lädermann, A., Nové-Josserand, L., & Smithers, C. J. (2012). Do the indications, results, and complications of reverse shoulder arthroplasty change with surgeon's experience? *Journal of Shoulder and Elbow Surgery*, *21*(11), 1470–1477. https://doi.org/10.1016/j.jse.2011.11.010

Wierks, C., Skolasky, R., Hun Ji, J., & McFarland, E. (2009). Reverse Total Shoulder Replacement: Intraoperative and Early Postoperative Complications. *Clin Orthop Relat Res*, *467*, 225–234.

Yoon, S., Park, T., Lee, J., & Kim, J. (2019). A Study on Transfer Effectiveness and Appropriate Training Hours in Airplanes Simulators. *ITEC Proceedings*.