Virtual Reality Simulation for Shoulder Arthroscopy: A Three-Year Follow-up Study of Individual Skill Progression

George Pappas, MD, PhD, Andreas H. Gomoll, MD, Brian Forsythe, MD, and Jon J. P. Warner, MD
Harvard Shoulder Service, Brigham and Women's Hospital and Massachusetts General Hospital

ABSTRACT

Previous studies have demonstrated a correlation between surgical experience and performance on a virtual reality arthroscopy simulator but only provided single time point evaluations. Additional longitudinal studies are necessary to confirm the validity of virtual reality simulation before these teaching aids can be more fully recommended for surgical education. We hypothesized that subjects would demonstrate improved performance on shoulder arthroscopy simulator retesting several years after an initial baseline evaluation, commensurate with their advanced surgical experience. After gaining further arthroscopic experience, 10 orthopaedic residents underwent retesting three years after initial evaluation on a Procedicus virtual reality shoulder arthroscopy simulator. Using a paired *t* test, simulator parameters were compared in each subject before and after additional arthroscopic experience. Subjects were evaluated for time to completion, number of probe collisions with the tissues, average probe velocity, and distance traveled with the tip of the simulated probe compared to an optimal computer-determined distance. In addition, to evaluate consistency of simulator performance, results were compared to historical controls of equal experience. Subjects significantly improved their performance on simulator retesting three years after initial evaluation. Scores improved significantly ($P < 0.02$ for all) in the 4 simulator parameters: completion time (–51%), probe collisions (–29%), average velocity, (122%), and distance traveled (–32%). With the exception of probe velocity, there were no significant differences between the performance of this group and that of a historical group with equal experience, suggesting that groups with similar arthroscopic experience demonstrate equivalent scores on the simulator. Thus it may eventually be possible to establish simulator benchmarks to indicate likely arthroscopic skill. These results further validate the use of surgical simulation as an important tool for the evaluation of surgical skills.

George Pappas, MD, Resident, Harvard Combined Orthopaedic Program
Andreas Gomoll, MD, Clinical Instructor of Orthopaedic Surgery
Brian Forsythe, MD, Former Resident, Harvard Combined Orthopaedic Program
Jon JP Warner MD, Chief, Partners Shoulder Service; Professor of Orthopaedic Surgery

Address Correspondence to:

Department of Orthopedic Surgery
Brigham and Women's Hospital
75 Francis Street, PBB A-2
Boston, MA 02115

INTRODUCTION

The rapid evolution of surgical techniques and procedures is challenging the established methods of surgical education. Several factors have prompted a re-evaluation of the traditional one-on-one apprenticeship model of education: work-hour restrictions reduce surgical experience for residents; cost pressures[1,3] and concerns for patient safety[2] present disincentives for resident involvement in the operating room; and increasingly complex and minimally invasive procedures are associated with steep learning curves that cannot be adequately mastered via current operating room education alone. Several educational and professional bodies, including the American Academy of Orthopaedic Surgeons and the American Board of Orthopaedic Surgery, have recognized the growing need for alternative methods of education and are pursuing venues such as cadaveric laboratories, anatomical models, and, more recently, virtual reality (VR) computer simulation.[2,6,8] However, the widespread introduction of VR simulators for surgical education has been hampered by concerns that simulators would be more likely to evaluate a subject's aptitude for video games rather than their surgical skills.[5,9] These concerns were initially addressed by studies,[7,10,11] including our own,[4] that demonstrated a clear correlation between surgical experience and performance on a VR arthroscopy simulator. These studies validated the use of VR simulators for the evaluation of surgical skills, however they were limited in that they provided only a single time point evaluation of this correlation. The next step in the development of more advanced simulator applications for surgical education would involve prospective investigations demonstrating that improvement in surgical experience correlated with improvement in simulator performance.[6] Thus, we retested subjects who had participated in our initial validation study[4] and who had demonstrated significant differences in simulator performance between four groups of different levels of surgical experience.

We hypothesized that subjects who had gained arthroscopic experience would demonstrate improved simulator performance on retesting three years after baseline evaluation. Confirming this hypothesis would support test responsiveness by providing a correlation between changing surgical experience and changing simulator performance. Furthermore, we hypothesized that subjects' current performance level would be commensurate with their advanced experience level, as compared with a historical group of equal experience level.

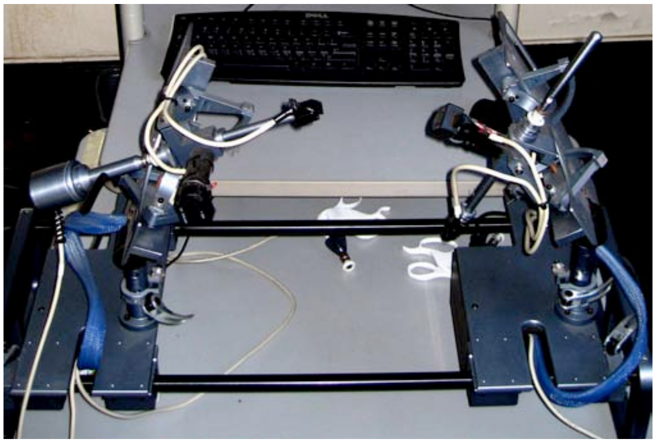

Figure 1: Actuator controls for Mentis Procedicus simulator (arthroscope on left, probe on right).

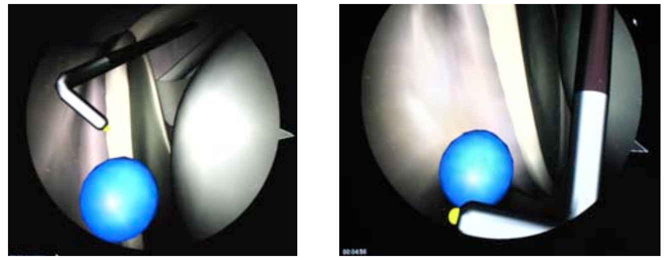

Figure 3: Manipulation test. Subject must touch tip of probe to blue sphere, which appears at random locations within the simulated glenohumeral joint.

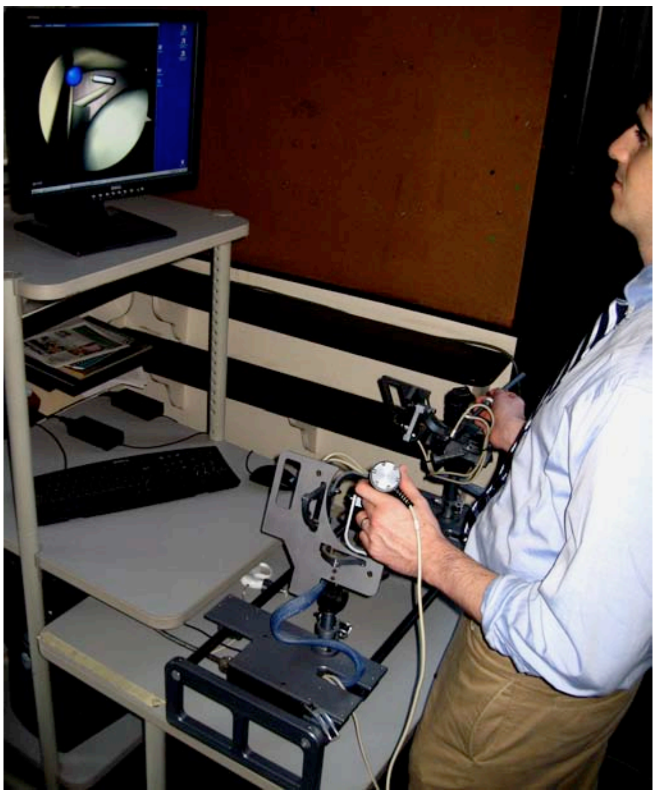

Figure 2: Test subject using Procedicus shoulder arthroscopy simulator.

MATERIALS AND METHODS

In our original study, we included 43 subjects: 8 with no arthroscopic experience (medical students and surgical interns), 11 with limited experience (second- and third-year orthopaedic residents), 14 with moderate experience (fourth and fifth-year residents), and 10 with extensive experience (sports medicine fellows and attendings). We focus here on the 10 subjects who initially had no experience (5 subjects) or limited experience (5 subjects) with shoulder arthroscopy, who were still enrolled at our residency program and therefore available for retesting 3 years after the initial evaluation. At the time of the initial testing, subjects with limited arthroscopy experience had participated, on average, in 20 shoulder arthroscopic surgeries. On re-evaluation, all subjects had completed 3 additional rotations in sports medicine (each of 4 weeks duration) and had participated, on average, in 60 shoulder arthroscopic surgeries.

As previously described,[4] participants were evaluated with a standardized protocol on a Procedicus arthroscopy simulator (Mentice Corp, Göteborg, Sweden). Each subject performed 6 repetitions of the same training module using the actuators to control a simulated arthroscope and probe (Figures 1 and 2). The module required the participant to locate and probe a simulated target (blue ball) within the shoulder joint (Figure 3). Successful contact between a simulated arthroscopic probe and the target caused the latter to disappear and be replaced by another target in a different, random location. This process was repeated 10 times (11 targets); thereafter the simulator calculated the following 4 parameters: time to completion of the module; number of probe collisions with the tissues; average velocity of probe movement; and the distance traveled with the tip of the simulated probe compared with a computer-determined optimal distance. For our primary analysis, we used paired *t* tests to compare the subjects' performance on each of the simulator parameters at study baseline versus after 3 years of followup. For additional analysis, to evaluate the consistency of the simulator assessment, we compared the current performance of these subjects, who now had moderate experience, with the performance of another subject group of moderate experience from our previous publication (historic controls, N = 14). In all statistical tests, the level of significance was set at $P < .05$.

RESULTS

Among subjects who had gained arthroscopic experience during the 3 years of follow-up, we found that simulator performance had improved from baseline to re-evaluation on all 4 parameters (Table 1). We also compared simulator performance of the current group of 10 subjects with moderate arthroscopic experience to the subject group with moderate experience from our previous publication. Simulator performance was very similar in the two study groups: there were no statistically significant differences for time to completion, hook collisions,

	Baseline Scores		Current Scores				Historical Control Group		
	Average	SD	Average	SD	% Change	P Value	Average	SD	P Value [b]
Time to completion (sec)	101.4	33.0	49.3	13.9	-51	< .001	59.9	14.4	0.1
Hook collisions (n)	34.7	15.0	24.6	12.5	-29	0.025	21.2	3.7	0.4
Probe velocity (cm/sec)	0.22	0.09	0.49	0.11	122	< .001	0.33	0.08	0.001
Actual more than optimal path (n)	7.7	2.3	5.2	1.9	-32	0.003	4.9	0.9	0.5

TABLE 1. Performance Scores of the Same Group and Historical Control Group[a]

[a] Historical control group was comprised of subjects with similar experience level. SD, standard deviation
[b] With respect to "current" test scores.

and distance traveled; only one parameter, velocity, was significantly different compared with historical controls (Table 1).

DISCUSSION

This study investigated the performance of subjects who were retested 3 years after initial evaluation on a VR simulator for shoulder arthroscopy. During baseline evaluation, these subjects possessed either limited or no experience with arthroscopy. By the time of re-evaluation, the subjects had advanced to moderate experience level. Comparing current and baseline simulator performance, subjects improved significantly in all tested parameters. Moreover, there were no significant differences in most of the parameters when comparing the current performance with a historical group of the same (moderate) experience level. This provides preliminary evidence that the simulator yields consistent results across groups with similar surgical experience.

This investigation reports individual longitudinal changes in simulator performance on retesting three years after an initial evaluation. The correlation between simulator skills and surgical experience reported in earlier studies helped demonstrate that simulators were evaluating surgical skills rather than video gaming skills and basic hand-eye coordination.[4,7,11] Our previous study had demonstrated differences between independent groups of limited or no experience versus moderate arthroscopic experience of 41% in time to completion, of 39% in hook collisions, of 52% in velocity, and of 37% in distance traveled. This new study suggests gains in experience over time within the same group are accompanied by gains in simulator performance. Interestingly, the extent of improvement in simulator performance was consistent with the differences we found across groups in our previous study: rising from a level of no or limited experience to one of moderate experience improved time to completion in our current subjects by 51%, hook collisions decreased by 29%, velocity increased by 122%, and distance traveled decreased by 32%. These data provide important additional support for the intrinsic value of simulator testing by showing substantial individual improvement in simulator performance when additional surgical experience was gained. Also, these results demonstrate internal consistency, since groups with comparable differences in surgical experience showed consistent differences in simulator performance, irrespective of whether the differences in experience were between two different groups or within the same group tested several years apart.

Comparing the simulator performance of current subjects with a historical control group of similar experience level demonstrated no significant differences in 3 of the 4 parameters tested, except for probe velocity, which was significantly higher in current subjects. This could represent the increased familiarity of our current subjects with the simulator, since they had the opportunity to undergo simulator testing twice, compared with only once for the historical control. Interestingly, two medical students who were retested but not included in this study because they had not gained any actual arthroscopic experience also demonstrated significantly faster probe velocity as the only difference to their original performance; all other parameters remained essentially unchanged. While it is impossible to base any conclusion on a group of only two subjects, this finding warrants further investigations with a larger number of subjects.

Some study limitations should be considered. Our study included a relatively small sample of subjects who were retested after gaining additional arthroscopic experience. Our original study included 19 participants with no or limited arthroscopic experience; since medical students and residents are a geographically highly mobile group, we were able to re-evaluate only 10 of these subjects 3 years later. However, there was no reason to believe that the 10 subjects we retested differed meaningfully from the 9 subjects we could not retest, who were either orthopaedic residents that had graduated or medical students who matched at other orthopaedic programs. Furthermore, the consistency in simulator performance of our current subjects with moderate arthroscopic experience was very similar to the performance of historic controls with moderate experience, which is additional evidence of the validity of our findings. Nonetheless, additional studies with larger sample sizes are warranted to further investigate simulator evaluation and training.

CONCLUSION

Individuals who had gained surgical experience in the interval between two identical tests with a VR arthroscopy simulator demonstrated substantially improved results on the simulator as well, further validating the use of VR simulation for the evaluation of surgical skills. Additionally, across independent groups with equivalent surgical experience, similar performance can be expected on the simulator parameters, suggesting it may eventually be possible to establish simulator benchmarks to indicate likely arthroscopic skill. This study and previous studies provide a foundation for additional research to assess the role of VR simulation not only as a tool for the evaluation but also the teaching of surgical skills, predominantly hand-eye coordination and surgical anatomy. More complex tasks, such as the teaching of actual surgical procedures, for example, rotator cuff or instability repair, will require more advanced simulators, which are currently under development.

References

1. **Bridges M, Diamond DL.** The financial impact of teaching surgical residents in the operating room. *Am J Surg.* 1999;177(1):28-32.
2. **Cannon WD, Eckhoff DG, Garrett WE Jr, Hunter RE, Sweeney HJ.** Report of a group developing a virtual reality simulator for arthroscopic surgery of the knee joint. *Clin Orthop Relat Res.* 2006;442:21-29.
3. **Farnworth LR, Lemay DE, Wooldridge T, et al.** A comparison of operative times in arthroscopic ACL reconstruction between orthopaedic faculty and residents: the financial impact of orthopaedic surgical training in the operating room. *Iowa Orthop J*. 2001;21:31-35.
4. **Gomoll AH, O'Toole RV, Czarnecki J, Warner JJ.** Surgical experience correlates with performance on a virtual reality simulator for shoulder arthroscopy. *Am J Sports Med.* 2007;35(6):883-888.
5. **Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J.** Impact of hand dominance, gender, and experience with computer games on performance in virtual reality laparoscopy. *Surg Endosc*. 2003;17(7): 1082-1085.
6. **Michelson JD.** Simulation in orthopaedic education: an overview of theory and practice. *J Bone Joint Surg Am*. 2006;88(6):1405-1411.
7. **Pedowitz RA, Esch J, Snyder S.** Evaluation of a virtual reality simulator for arthroscopy skills development. *Arthroscopy*. 2002;18(6):E29.
8. **Poss R, Mabrey JD, Gillogly SD, et al.** Development of a virtual reality arthroscopic knee simulator. *J Bone Joint Surg Am*. 2000;82(10): 1495-1499.
9. **Rosenberg BH, Landsittel D, Averch TD.** Can video games be used to predict or improve laparoscopic skills? *J Endourol*. 2005;19(3):372-376.
10. **Smith S, Wan A, Taffinder N, Read S, Emery R, Darzi A.** Early experience and validation work with Procedicus VA—the Prosolvia virtual reality shoulder arthroscopy trainer. *Stud Health Technol Inform*. 1999;62:337-343.
11. **Srivastava S, Youngblood PL, Rawn C, Hariri S, Heinrichs WL, Ladd AL.** Initial evaluation of a shoulder arthroscopy simulator: establishing construct validity. *J Shoulder Elbow Surg*. 2004;13(2):196-205.