

# Impact of Simulation Training on Diagnostic Arthroscopy Performance: A Randomized Controlled Trial



Kevin C. Wang, M.D., Eamon D. Bernardoni, M.D., Eric J. Cotter, M.D., Brian J. Cole, M.D., M.B.A., Nikhil N. Verma, M.D., Anthony A. Romeo, M.D., Charles A. Bush-Joseph, M.D., Bernard R. Bach, M.D., and Rachel M. Frank, M.D.

**Purpose:** To determine the impact of training on a virtual reality arthroscopy simulator on both simulator and cadaveric performance in novice trainees. **Methods:** A randomized controlled trial of 28 participants without prior arthroscopic experience was conducted. All participants received a demonstration of how to use the ArthroVision Virtual Reality Simulator and were then randomized to receive either no training (control group,  $n = 14$ ) or a fixed protocol of simulation training ( $n = 14$ ). All participants took a pretest on the simulator, completing 9 tasks ranging from camera-steadying tasks to probing structures. The training group then trained on the simulator (1 time per week for 3 weeks). At week 4, all participants completed a 2-part post-test, including (1) performing all tasks on the simulator and (2) performing a diagnostic arthroscopy on a cadaveric knee and shoulder. An independent, blinded observer assessed the performance on diagnostic arthroscopy using the Arthroscopic Surgical Skill Evaluation Tool scale. To compare differences between non-normally distributed groups, the Mann-Whitney  $U$  test was used. An independent-samples  $t$  test was used for normally distributed groups. The Friedman test with pair-wise comparisons using Bonferroni correction was used to compare scores within groups at multiple time points. Bonferroni adjustment was applied as a multiplier to the  $P$  value; thus, the  $\alpha$  level remained consistent. Significance was defined as  $P < .05$ . **Results:** In both groups, all tasks except task 5 (in which completion time was relatively fixed) showed a significant degree of correlation between task completion time and other task-specific metrics. A significant difference between the trained and control groups was found for post-test task completion time scores for all tasks. Qualitative analysis of box plots showed minimal change after 3 trials for most tasks in the training group. There was no statistical correlation between the performance on diagnostic arthroscopy on either the knee or shoulder and simulation training, with no difference in Arthroscopic Surgical Skill Evaluation Tool scores in the training group compared with controls. **Conclusions:** Our study suggests that an early ceiling effect is shown on the evaluated arthroscopic simulator model and that additional training past the point of proficiency on modern arthroscopic simulator models does not provide additional transferable benefits on a cadaveric model. **Level of Evidence:** Level I, randomized controlled trial.

Graduate medical education, specifically surgical education, has traditionally been rooted in an apprenticeship model. The “see one, do one, teach one” approach has guided education in the operating room

and has led to the development of surgical education as it stands today. However, there are rising concerns regarding burnout—a condition correlated with longer work hours, a skewed work-life balance, and

*From the Department of Orthopaedic Surgery, Mount Sinai Hospital (K.C.W.), New York, New York; Department of Orthopaedic Surgery, University of Wisconsin (E.D.B., E.J.C.), Madison, Wisconsin; Department of Orthopaedic Surgery, Rush University Medical Center (B.J.C., N.N.V., A.A.R., C.A.B.-J., B.R.B.), Chicago, Illinois; and Department of Orthopaedic Surgery, University of Colorado (R.M.F.), Boulder, Colorado, U.S.A.*

*The authors report the following potential conflicts of interest or sources of funding: This study was funded by an American Board of Orthopaedic Surgery education grant. The funding source had no input into study development, design, or data analysis. B.J.C. receives research support from Aesculap/B.Braun, Geistlich, National Institutes of Health, and sanofi-aventis; is on the editorial or governing board of the American Journal of Orthopedics, American Journal of Sports Medicine, Arthroscopy, Cartilage,*

*Journal of Bone and Joint Surgery—American, and Journal of Shoulder and Elbow Surgery; receives stock or stock options from Aqua Boom, Biomerix, Giteliscopes, and Ossio; receives IP royalties and research support from and is a paid consultant for Arthrex; is a board or committee member of the Arthroscopy Association of North America and International Cartilage Repair Society; receives other financial or material support from Athletico, JRF Ortho, and Tornier; receives IP royalties from DJ Orthopaedics and Elsevier Publishing; is a paid consultant for Flexion; receives publishing royalties and financial or material support from Operative Techniques in Sports Medicine; is a paid consultant for and receives stock or stock options from Regentis; receives other financial or material support from and is a paid consultant for Smith & Nephew; and is a paid consultant for and receives research support from Zimmer. N.N.V. is a board or committee member of*

potentially negative patient outcomes—among orthopaedic surgery trainees.<sup>1-3</sup> Given these concerns, the traditional method of technical training—emphasizing repetition and exposure—has come under scrutiny as educators attempt to improve on this framework. These changing attitudes have been reflected in recent changes in work-hour restrictions and formal surgical skill training programs put forth by the Accreditation Council for Graduate Medical Education.

Changing attitudes toward surgical education have developed in parallel with new educational tools. The development of modern computers and virtual reality systems offers exciting opportunities to augment surgical training. Specifically, these technologies enable development of technical skills in a controlled, low-risk environment that allows trainees to break down complex techniques into digestible components that can be practiced to the point of competency. New training models enable mastery of these component skills to allow trainees to perform complex tasks more effectively,<sup>4</sup> and these models allow measurement of stepwise procedural proficiency. Investigations showing the efficacy of proficiency-based progression in the attainment of scope-based skills (arthroscopy and laparoscopy) support this theory by showing improved performance of trainees in proficiency-based progression programs compared with controls.<sup>5,6</sup>

Previous investigations have generally supported some degree of skill transfer from practice on arthroscopic simulator models to performance on either cadaveric models or live arthroscopy.<sup>7-18</sup> However, some investigations have challenged the idea that simulator training provides generalizable skill acquisition.<sup>19,20</sup> In addition, the questions of how much simulator training is necessary and if there is an upper limit to the benefits of simulator training hold particular importance as new simulator technologies are being

integrated into residency training programs. The purpose of this study was to determine the impact of training on a virtual reality arthroscopy simulator on both simulator and cadaveric performance in novice trainees. We hypothesized that training on a simulator would result in improvements in arthroscopy performance on a cadaveric specimen and secondarily hypothesized that training on the simulator would show a ceiling effect.

## Methods

### Participants

After institutional review board approval was received (No. 16082301), a total of 28 novice trainees (preclinical medical and premedical students) at a single institution were invited to participate. All volunteers gave informed consent and were recruited on a voluntary basis without compensation. Participation status did not influence academic standing. Subjects were enrolled between November and December 2016. Initial testing and training were conducted from December 2016 to January 2017, and post-test data collection occurred in February 2017. Subjects were excluded if they had any previous arthroscopy experience or previous formal arthroscopy simulator training. Demographic information was collected prior to testing and included subject age, sex, video game use, and handedness using subject-reported surveys; video game use was collected on a 5-point Likert scale. There were no changes to the inclusion criteria after trial commencement.

### Study Design

The study was a single-blinded, prospective, randomized controlled trial with a parallel-group design. All subjects underwent an initial simulator pretest on an ArthroVision Virtual Reality Simulator (Swemac,

---

*American Shoulder and Elbow Surgeons, American Orthopaedic Society for Sports Medicine, and Arthroscopy Association of North America; is a paid consultant for and receives research support from Arthrex; is on the editorial or governing board of and receives publishing royalties and financial or material support from Arthroscopy; receives research support from Arthro-surface, DJ Orthopaedics, Össur, Athletico, ConMed Linvatec, Miomed, and Mitek; receives stock or stock options from Cymedica and Omeros; is on the editorial or governing board of Journal of Knee Surgery and SLACK; is a paid consultant for and receives stock or stock options from Minivasive; is a paid consultant for Orthospace; receives IP royalties and research support from Smith & Nephew; and receives publishing royalties and financial or material support from Vindico Medical-Orthopedics Hyperguide. A.A.R. receives research support from Aesculap/B.Braun, Histogenics, NuTech, Medipost, Smith & Nephew, Zimmer, and Orthospace; is on the editorial or governing board of the American Shoulder and Elbow Surgeons, Orthopedics, SAGE, and Wolters Kluwer Health—Lippincott Williams & Wilkins; receives other financial or material support and research support from and is a paid consultant and paid presenter or speaker for Arthrex; receives other financial or material support from the Arthroscopy Association of North America and MLB; is a board or committee member of Atrion Orthopaedics; is a board or*

---

*committee member and is on the editorial or governing board of Orthopedics Today; receives publishing royalties and financial or material support from Saunders/Mosby-Elsevier; and is on the editorial or governing board of and receives publishing royalties and financial or material support from SLACK. C.A.B.-J. is on the editorial or governing board of the American Journal of Sports Medicine; is a board or committee member of the American Orthopaedic Society for Sports Medicine; and receives stock or stock options in Cresco Lab. B.R.B. receives research support from Arthrex, ConMed Linvatec, DJ Orthopaedics, Össur, Smith & Nephew, and Tornier and publishing royalties and financial or material support from SLACK. Full ICMJE author disclosure forms are available for this article online, as [supplementary material](#).*

*Received February 7, 2019; accepted July 9, 2019.*

*Address correspondence to Kevin C. Wang, M.D., One Gustave Levy Place, Box 1188, New York, NY 10029, U.S.A. E-mail: [wangkevinc@gmail.com](mailto:wangkevinc@gmail.com)*

*© 2019 by the Arthroscopy Association of North America. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).*

*2666-061X/1997*

*<https://dx.doi.org/10.1016/j.asmr.2019.07.002>*

Linköping, Sweden). Subjects were then allocated to either a simulator training group ( $n = 14$ ) or no training group (control group,  $n = 14$ ) by the primary author (K.C.W.) using a computer-based random number generator set to create 2 equal-sized groups. Sample size was limited by the number of volunteers and is comparable with previously published trials, but no power analysis was conducted.<sup>6,16</sup> Allocation was concealed from participants and investigators until after completion of the pretest by the computer.

After completion of the training curriculum, all subjects from both groups participated in an additional, identical simulator session (simulator post-test). After this post-test, all subjects were shown two 5-minute videos of the appropriate technique for diagnostic arthroscopy of each joint and given a 5-minute overview of relevant anatomy prior to performing a 5-minute diagnostic arthroscopy on a cadaveric knee and shoulder. Simulation training and post-test data collection were conducted at a central location with no access to the simulator between sessions.

### Simulator Training

The ArthroVision Virtual Reality Simulator was used for training and testing (Fig 1). Each identical simulator session consisted of 9 tasks. A degree of construct validity on this simulator model has been previously established through the relation to other variables; this simulator model showed discriminative ability based on surgical level of training.<sup>21</sup> Session scores were obtained from the default metrics provided by the simulator. The tasks and measurement metrics collected automatically by the simulator program are shown in Table 1. To compare between trials and subjects, completion time was used because it was the only metric reported consistently between tasks.

The training group performed a simulator session once weekly for 3 consecutive weeks. Each simulator session consisted of a single run through each of the 9 tasks. The control group had no access to the simulator. Both groups completed a post-test simulator session after the training period was completed.

### Cadaveric Post-Test

After the training period, all subjects completed a cadaveric post-test on both a fresh-frozen human knee and shoulder. Subjects were shown instructional videos depicting the steps of diagnostic arthroscopy and were provided a diagram of anatomic landmarks. Basic information on the assessment criteria and a checklist of tasks were provided during the instructional period and while subjects were performing the procedure (Appendix 1). All videos were prerecorded and narrated by the senior author (R.M.F.). The primary

outcome collected from the cadaveric post-test was the Arthroscopic Surgical Skill Evaluation Tool (ASSET) score, a global rating scale for arthroscopic skill.<sup>22</sup> All procedures were recorded using the scope camera, blinded by removing any subject-identifying features, and graded using the ASSET by a fellowship-trained orthopaedic sports surgeon who was blinded to group allocation (R.M.F.). There were no changes in trial outcomes after the trial was commenced. After cadaveric post-test collection, the trial was completed.

### Statistical Analysis

All statistical analysis was performed using SPSS software (IBM, Chicago, IL). The Shapiro-Wilk test for normality was used to assess the normality of the distribution of variables of interest. For non-normally distributed variables, the Mann-Whitney  $U$  test was applied, and for normally distributed points, an independent-samples  $t$  test was used. To compare nominal variables, the Fisher exact test was used. The Wilcoxon signed rank test was used to compare pretest scores with post-test scores within groups. The Friedman test was run to determine if there were differences in scores during the pretest, post-test, and training period in the simulator-trained group. Pair-wise comparisons were performed with Bonferroni correction for multiple comparisons. In the SPSS statistical package, this is performed by multiplying the raw  $P$  value by the appropriate Bonferroni adjustment rather than adjusting the  $\alpha$  level. To assess the relation between test scores and subject demographic traits, a Spearman rank order correlation was run. An independent-samples  $t$  test was used to compare the performance on final cadaveric testing between the simulator-trained and control groups.

### Results

Twenty-eight participants completed the study, with 14 participants in each group. Except for 1 subject in the control group who did not complete the cadaveric post-test because of scheduling conflicts, all subjects completed all portions of the trial (Fig 2). There was no significant difference in age (mean  $\pm$  standard deviation,  $24.86 \pm 2.38$  years in control group vs  $24.36 \pm 2.59$  years in simulator group;  $P = .599$ ), sex (71.4% male subjects in both groups,  $P = .999$ ), right-hand dominance (92.9% in control group vs 85.7% in simulator group,  $P = .999$ ), or video game use (mean  $\pm$  standard deviation,  $2.61 \pm 0.88$  in control group vs  $2.13 \pm 0.82$  in simulator group;  $P = .146$ ) between the groups. In both groups, all tasks except task 5 (in which completion time was fixed) showed a significant degree of correlation between task completion time and other computer-collected, task-specific metrics (Table 2).



**Fig 1.** ArthroVision simulator setup. All simulations used the default settings and manufacturer-determined computerized scoring metrics.

A significant difference in pretest completion time was found for task 1 ( $P = .044$ ) between the simulator and control groups, favoring the simulator group; otherwise, no significant differences in completion time were found between the 2 groups (Table 3). Pretest completion time was not significantly correlated with video game experience ( $P = .152$ ).

A significant difference between the simulator and control groups was noted for post-test completion time for all measured tasks except task 5 (Table 4). Post-test

completion times for both groups were significantly different from the respective pretests. Qualitative analysis of box plots showed minimal change after 3 trials for most tasks in the training group (Fig 3). The Friedman test also showed no significant change between the final 3 trials (2 training sessions and the post-test) for tasks 3, 4, 6, and 8. The Bonferroni-adjusted  $P$  value for each of these tasks was .199, .292, .058, and .232, respectively, with an  $\alpha$  level of .05. For all other tasks, pair-wise comparisons found no

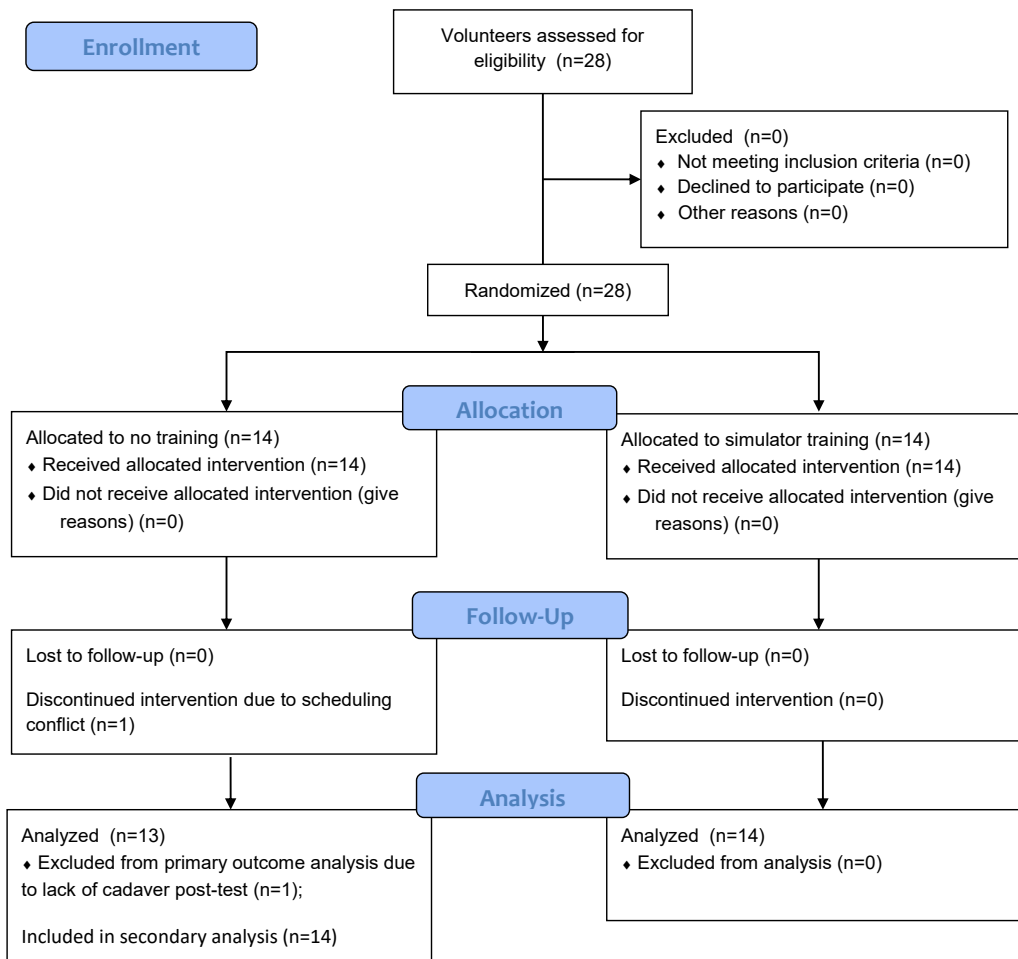
**Table 1.** Simulator Tasks and Metrics

Task	Description	Metrics Recorded by Simulator
1	Steadying camera and telescoping on a target	Time, path, path in focus
2	Performing periscopic movements around targets in a circle	Time, path
3	Tracking a moving target	Time, time out of focus, distance deviation, centering deviation
4	Achieving deliberate linear scope motion with the camera	Time, path
5	Tracking and probing a moving target	Time out of focus, manipulating time out of focus, time touching track, distance and centering deviation
6	Performing periscopic movements around a target	Time, path, telescoping path, XY path, view direction deviation
7	Measuring with a probe	Time, path, size deviation
8	Steadying camera, telescoping, and probing	Time, time out of focus, scope and probe path, manipulating time out of focus
9	Steadying camera, telescoping, and probing different position	Time, time out of focus, scope and probe path, manipulating time out of focus

significant difference in completion time between trials 4 and 5 (last training session and final post-test session).

A negative correlation was found between completion time and post-test cadaveric ASSET score (Spearman  $\rho = -0.084$  for pretest and Spearman  $\rho = -0.094$  for post-test); however, this did not reach significance ( $P = .666$  for pretest and  $P = .649$  for post-test). The

primary outcome measurement, cadaveric post-test ASSET score, was  $19.25 \pm 2.46$  in the simulator group and  $18.00 \pm 7.43$  in the control group. These scores were not significantly different between groups ( $P = .555$ ). This finding showed a Cohen  $d$  effect size of 0.23 (95% confidence interval,  $-0.528$  to  $0.987$ ), showing a small effect of simulator training on cadaveric post-test



**Fig 2.** CONSORT (Consolidated Standards of Reporting Trials) flow diagram showing patient recruitment, follow-up, and analysis.

**Table 2.** Spearman  $\rho$  Correlation for Task Completion and All Other Simulator Scoring Metrics

	CE/P Value								
	Task 1	Task 2	Task 3	Task 4	Task 6	Task 7	Task 8	Task 9	
Path	0.749/.002	0.974/<.001	0.851/<.001	0.807/<.001	0.811/<.001	0.780/.001	—	—	0.895/<.001
Path in focus	0.798/.001	—	—	—	—	—	—	—	—
Telescoping path	—	—	0.820/<.001	—	—	—	—	—	—
XY path	—	—	0.833/<.001	—	—	—	—	—	—
View direction deviation	—	—	—	—	—	—	—	—	—
Total distance deviation	—	—	—	—	—	—	—	—	—
Total centering deviation	—	—	—	—	—	—	—	—	—
Size error 1*	—	—	—	—	0.059/.84	—	—	—	—
Probe path	—	—	—	—	—	0.851/<.001	—	—	0.846/<.001
Time out of focus	—	—	—	—	—	0.925/<.001	—	—	0.912/<.001
Manipulation time out of focus	—	—	—	—	—	0.53/.051	—	—	0.714/.004
Total track touch time	—	—	—	—	—	—	0.262/.366	—	—

CE, correlation coefficient.

\*Representative of 5 size errors (1 for each object measured during the task).

performance. The Levene test for equality of variances did not show any significant difference in variance between groups ( $F = 2.023, P = .167$ ). No harm or unintended consequences were reported for either group.

### Discussion

The principal findings of this study suggest that (1) there is a significant improvement from pretest to post-test completion time on the simulator model in both groups, (2) a ceiling effect exists in simulator training that can be rapidly achieved, (3) additional training after reaching the point of the ceiling effect does not have a significant impact on performance on diagnostic arthroscopy, and (4) improved performance on simulator testing did not correlate significantly with the ASSET score on cadaveric arthroscopy. Our investigation showed that there was a significant improvement from pretest to post-test completion time on the simulator model in both groups. The training group showed significantly more improvement than the control group from simulator pretest to simulator post-test. However, the training group showed a ceiling effect after 3 sessions with the simulator for most tasks, with no significant differences found between trials 4 and 5 for all tasks; most of the gains in performance occurred between the initial exposure to the simulator (pretest) and the second exposure to the simulator, possibly because of the low complexity of the simulated task. It was also noted that a single session on the simulator enabled a rapid acquisition of skills that persisted for at least 1 month (control group). In summary, we have shown that simulator training results in improved simulator performance but does not significantly improve performance in diagnostic arthroscopy.

Previous investigations of simulator training have shown transfer validity of skills between simulator models,<sup>7,17</sup> cadaveric models,<sup>6,8,11,12,14,15,17</sup> and live arthroscopy,<sup>10,13,16</sup> whereas 2 studies found no direct transfer validity between simulator models.<sup>19,20</sup> A recent systematic review and meta-analysis of studies investigating the impact of simulator training on performance showed a strong effect on simulator performance and a moderate effect on human-model (live or cadaveric) performance.<sup>18</sup> Camp et al.<sup>9</sup> showed significantly greater improvement in final, cadaveric post-test ASSET scores for cadaveric training compared with a similar amount of simulator training; however, simulator training was superior to no training. Butler et al.<sup>8</sup> reported diminishing marginal returns (less improvement with each subsequent trial) on a bench-top knee simulator. In their investigation, subjects continued to train until they no longer showed improvement on the simulator, showing that proficiency-based simulator training combined with cadaveric training resulted in fewer trials to achieve

**Table 3.** Median Task Completion Time on Simulator Pretest

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9
Median time (IQR), s									
Simulator group	123.7 (68.7)	130.4 (102.0)	138.7 (83.4)	26.3 (5.9)	73.6 (0.0)	212.9 (121.8)	321.4 (128.7)	173.6 (102.6)	133.5 (121.6)
Control group	169.7 (128.0)	104.9 (104.4)	159.3 (183.2)	25.9 (7.0)	73.6 (0.0)	252.9 (125.6)	301.4 (224.6)	198.2 (140.4)	141.5 (235.2)
P value	.044	.427	.511	.867	.329	.265	.839	.946	.280

IQR, interquartile range.

proficiency on a cadaver than cadaveric training alone. Together, these investigations support the assertion that arthroscopy simulator training is transferable to arthroscopy on a human model by targeting the component skills required in arthroscopic surgery, such as bimanual dexterity and spatial awareness.

However, other investigations have challenged the generalizability of current simulators and the notion of skill transfer.<sup>19,20</sup> They support the argument that current simulators cannot yet reproduce the higher-level integration of many cognitive domains including anatomic knowledge, instrument handling, and anticipation of next steps required in real-life surgery. Ferguson et al.<sup>20</sup> showed diminishing marginal returns with additional training and examined the effects of over-training (continued practice after no observable improvement is shown) on 2 checklist-guided benchtop simulators, showing no transfer of skills between different simulator models based on completion time and score on a global rating scale. However, in this investigation, novice trainees may have improved their understanding of joint-specific anatomy and movement patterns with these simulators without developing more basic component skills. Ström et al.<sup>19</sup> used multiple, focused simulator models but only allowed subjects to train for a short duration on each model. These investigations raise the concern that transfer validity should not be assumed on simulator models: Device design, task applicability, context of training, and trainee integration of knowledge are all important contributors to skill transference, and current simulator models need further development.<sup>23</sup> The difference in novice trainee skill acquisition on the simulator (as determined by simulator completion time) versus the cadaver (as determined by ASSET score) in this investigation is likely attributable to incomplete skill transference owing to a combination of these factors.

Our study shows that simulator training to the point of diminishing marginal returns does not significantly improve performance on cadaveric arthroscopy. This finding supports the argument that the skills gained on these simulator models are not directly transferable to an operative setting. Simulator proficiency is rapidly acquired (in just a few hours) to the point at which the opportunity cost of additional training on these simulators (less time spent with cadavers or in the operating room) may not justify additional practice time. Previous literature has suggested that a proficiency-based simulation curriculum can provide improved performance over a traditional curriculum<sup>6,8</sup>; however, our data suggest that the time required to reach proficiency on a simulator is not extensive, and the return on further training on the simulator after this point may not translate into significant gains on human models. This should be

**Table 4.** Median Task Completion Time on Simulator Post-Test

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9
Median time (IQR), s									
Simulator group	60.5 (34.3)	53.1 (30.7)	64.3 (24.8)	20.4 (5.1)	73.6 (0.0)	117.6 (77.4)	139.7 (31.9)	74.2 (25.3)	49.8 (16.2)
Control group	92.4 (58.9)	77.9 (32.5)	107.1 (85.0)	25.3 (7.6)	73.6 (0.0)	196.3 (97.5)	255.4 (55.7)	116.1 (80.8)	85.3 (75.7)
P value	.01	.039	.014	.002	.650	.044	<.001	.001	<.001

IQR, interquartile range.

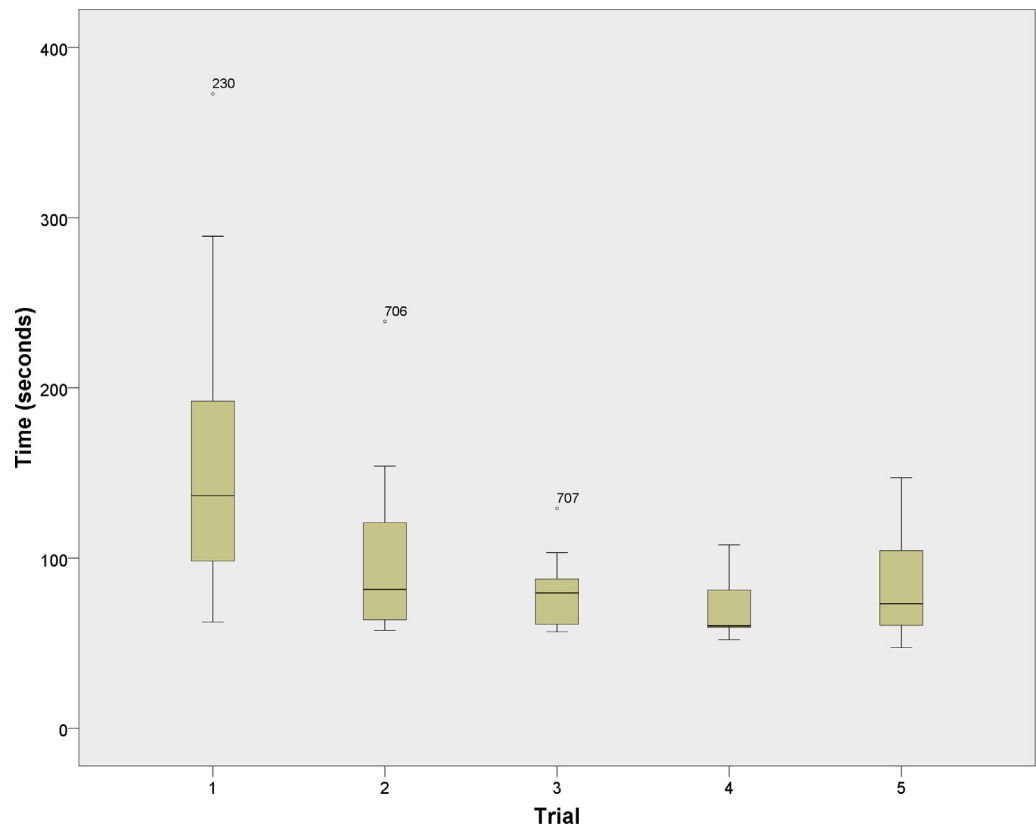
explored in future research and considered when designing simulation training curricula.

### Limitations

Our investigation was initially intended to provide insight into the impact of simulator training versus no training on cadaveric arthroscopy performance in a simulator-naive population. However, it faces several limitations. The sample size, as well as power, of this study was limited by the number of volunteers. This is a limitation faced by much of the previous literature in this area, and our sample size is comparable to that of other investigations. The lack of a true control group (simulator-naive subjects with no exposure to the simulator during the study) limits our interpretation of the impact of our simulator training program on subject performance on cadaveric arthroscopy. In this study, the differences in pretest and post-test simulator scores show that our pretest study design allowed for significant skill acquisition on the simulator, and thus, the simple act of performing a pretest on the simulator could be considered “training.” In fact, some studies have used a total simulator exposure time of 1 hour as a simulator training curriculum.<sup>11,16,19</sup> Thus, although it was not our intent on study design, this study may be better classified as a trial of “trained” versus “over-trained” subjects as opposed to “trained” versus “untrained.” Because of the lack of a true, simulator-naive control group, the interpretation of our results must be made with caution. Our study lacked a pretest session prior to cadaveric training because the additional use of cadavers was cost prohibitive in this investigation. Another limitation is that only a single senior author evaluated the ASSET scores of participants. The ASSET score has been previously validated to show high inter-rater reliability,<sup>22</sup> but inclusion of an additional rater would allow for score averaging. In addition, the properties of the simulator itself limited our analyses; there was only 1 consistent measurement between tasks. In the development of future simulator models, more consistency between tasks should be incorporated. Although our research was internally consistent, its external validity cannot be assumed. Our research was conducted on only a single simulator model, and our study population is not fully representative of the target population: novice orthopaedic trainees. Medical students were enrolled as novice trainees in this study in lieu of orthopaedic trainees to increase the number of potential volunteers. However, the limited baseline anatomic knowledge and technical training in our study population may limit the ability to show improvement; more advanced novices who have personally observed arthroscopy and have a solid understanding of fundamental anatomy may benefit more from simulator skill development.



**Fig 3.** Task completion time for simulator-trained group graphed by trial number. Trial 1 represents the initial pretest; trials 2, 3, and 4 are the additional training sessions; and trial 5 represents the simulator post-test.



## Conclusions

Our study suggests that an early ceiling effect is shown on the evaluated arthroscopic simulator model and that additional training past the point of proficiency on modern arthroscopic simulator models does not provide additional transferable benefits on a cadaveric model.

## References

- Barrack RL, Miller LS, Sotile WM, Sotile MO, Rubash HE. Effect of duty hour standards on burnout among orthopaedic surgery residents. *Clin Orthop Relat Res* 2006;449:134-137.
- Daniels AH, DePasse JM, Kamal RN. Orthopaedic surgeon burnout: Diagnosis, treatment, and prevention. *J Am Acad Orthop Surg* 2016;24:213-219.
- Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Stress and coping among orthopaedic surgery residents and faculty. *J Bone Joint Surg Am* 2004;86:1579-1586.
- Dougherty KM, Johnston JM. Overlearning, fluency, and automaticity. *Behav Anal* 1996;19:289-292.
- Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Ann Surg* 2002;236:458.
- Angelo RL, Ryu RKN, Pedowitz RA, et al. A proficiency-based progression training curriculum coupled with a model simulator results in the acquisition of a superior arthroscopic Bankart skill set. *Arthroscopy* 2015;31:1854-1871.
- Akhtar K, Sugand K, Wijendra A, et al. The transferability of generic minimally invasive surgical skills: Is there crossover of core skills between laparoscopy and arthroscopy? *J Surg Educ* 2016;73:329-338.
- Butler A, Olson T, Koehler R, Nicandri G. Do the skills acquired by novice surgeons using anatomic dry models transfer effectively to the task of diagnostic knee arthroscopy performed on cadaveric specimens? *J Bone Joint Surg Am* 2013;95:e15(1-8).
- Camp CL, Krych AJ, Stuart MJ, Regnier TD, Mills KM, Turner NS. Improving resident performance in knee arthroscopy: A prospective value assessment of simulators and cadaveric skills laboratories. *J Bone Joint Surg Am* 2016;98:220-225.
- Cannon WD, Garrett WE, Hunter RE, et al. Improving residency training in arthroscopic knee surgery with use of a virtual-reality simulator: A randomized blinded study. *J Bone Joint Surg Am* 2014;96:1798-1806.
- Dunn JC, Belmont PJ, Lanzi J, et al. Arthroscopic shoulder surgical simulation training curriculum: Transfer reliability and maintenance of skill over time. *J Surg Educ* 2015;72:1118-1123.
- Henn RF, Shah N, Warner JJP, Gomoll AH. Shoulder arthroscopy simulator training improves shoulder arthroscopy performance in a cadaveric model. *Arthroscopy* 2013;29:982-985.
- Howells NR, Gill HS, Carr AJ, Price AJ, Rees JL. Transferring simulated arthroscopic skills to the operating theatre: A randomised blinded study. *J Bone Joint Surg Br* 2008;90:494-499.
- Martin KD, Patterson D, Phisitkul P, Cameron KL, Femino J, Amendola A. Ankle arthroscopy simulation

- improves basic skills, anatomic recognition, and proficiency during diagnostic examination of residents in training. *Foot Ankle Int* 2015;36:827-835.
15. Rebolledo BJ, Hammann-Scala J, Leali A, Ranawat AS. Arthroscopy skills development with a surgical simulator: A comparative study in orthopaedic surgery residents. *Am J Sports Med* 2015;43:1526-1529.
  16. Waterman BR, Martin KD, Cameron KL, Owens BD, Belmont PJ. Simulation training improves surgical proficiency and safety during diagnostic shoulder arthroscopy performed by residents. *Orthopedics* 2016;39:e479-e485.
  17. Banaszek D, You D, Chang J, et al. Virtual reality compared with bench-top simulation in the acquisition of arthroscopic skill: A randomized controlled trial. *J Bone Joint Surg Am* 2017;99:e34.
  18. Frank RM, Wang KC, Davey A, et al. Utility of modern arthroscopic simulator training models: A meta-analysis and updated systematic review. *Arthroscopy* 2018;34:1650-1677.
  19. Ström P, Kjellin A, Hedman L, Wredmark T, Fellander-Tsai L. Training in tasks with different visual-spatial components does not improve virtual arthroscopy performance. *Surg Endosc* 2004;18:115-120.
  20. Ferguson J, Middleton R, Alvand A, Rees J. Newly acquired arthroscopic skills: Are they transferable during simulator training of other joints? *Knee Surg Sports Traumatol Arthrosc* 2017;25:608-615.
  21. Rose K, Pedowitz R. Fundamental arthroscopic skill differentiation with virtual reality simulation. *Arthroscopy* 2015;31:299-305.
  22. Koehler RJ, Goldblatt JP, Maloney MD, Voloshin I, Nicandri GT. Assessing diagnostic arthroscopy performance in the operating room using the Arthroscopic Surgery Skill Evaluation Tool (ASSET). *Arthroscopy* 2015;31:2314-2319.e2.
  23. Baldwin T, Ford K. Personnel psychology 1988. 41. *Pers Psychol* 1988;41:63-105.

## Appendix 1

### Basic Instructions

You will be assessed on the following:

- Safety: avoiding damage to structures (cartilage, ligaments, and so on)
- Field of view: adequately visualizing what you are examining (not zoomed too close or too far out)
- Camera dexterity: ability to keep the camera steady, centered, and correctly oriented
- Instrument dexterity: ability to maneuver instrument toward targets
- Bimanual dexterity: ability to coordinate movements with both hands
- Flow of procedure: ability to move from 1 step to the next
- Quality of procedure: completeness of the procedure
- Autonomy: ability to complete the procedure alone or with interventions
- Time: how long it takes you to get through all tasks on the checklist
- Area of focus: the number of times you need to look down at your hands rather than at the screen during the procedure

The proctors will only be able to read things off of the checklist for you (this will be available for you to read as well).

### Knee Arthroscopy Checklist

1. Evaluate the patella: Begin with the knee in extension; enter from the lateral side.
  - a. Rotate the lens to inspect the suprapatellar pouch.
  - b. Back out to inspect the undersurface of the patella.
  - c. Rotate the lens to inspect the lateral and medial patellar facets.
  - d. Bend the knee to check for patellar tracking in the trochlear groove.
  - e. Evaluate the patellofemoral articulation.
2. Evaluate the lateral gutter: Advance the arthroscope past the trochlea and patella, and rotate to view the lateral side of the knee.
  - a. Follow the lateral edge of the knee down to the lateral gutter.
  - b. Move the eyes of the arthroscope downward and medial in the lateral gutter to evaluate the popliteus tendon.
3. Evaluate the medial gutter: Advance the arthroscope back up the knee to the suprapatellar pouch, and continue medially into the medial gutter.
  - a. Follow the medial edge of the knee down to the medial gutter.
  - b. Move the eyes of the arthroscope downward to evaluate the medial gutter.
4. Evaluate the medial compartment: Return the arthroscope to the suprapatellar pouch, and back out along the trochlea to the notch.
  - a. Flex the knee while in the notch to open up the medial compartment.
  - b. Introduce the probe to evaluate the medial meniscus.
  - c. Be sure to gently probe above and below the meniscus to look for any tears.
  - d. Probe the medial femoral condyle articular cartilage and medial tibial plateau articular cartilage to assess for any damage.
5. Evaluate the cruciate ligaments: Return the arthroscope to the notch to visualize the anterior cruciate ligament and posterior cruciate ligament.
  - a. Use the probe to assess for integrity of the anterior cruciate ligament and posterior cruciate ligament.
6. Examine the lateral compartment: Move the knee into a figure-of-4 position to widen the lateral compartment, and move the arthroscope into the lateral compartment.
  - a. Evaluate the lateral compartment of the knee: Probe above and below the meniscus.
  - b. Probe the lateral femoral condyle and lateral tibial plateau to evaluate for cartilage damage.

### Shoulder Arthroscopy Checklist (Beach Chair)

Note that because the shoulder is smaller than the knee, most of the movement in the shoulder will be with the “eyes” of the arthroscope.

1. Establish the locations of the glenoid (socket) and the humeral head (ball).
2. Examine the long head of the biceps tendon
  - a. Probe the long head of the biceps, and pull it into the joint to examine it.
3. Examine and probe the superior labrum where the biceps inserts.
4. Inspect and probe the posterior labrum.
5. Inspect the inferior pouch of the humerus.
6. Inspect and probe the glenoid articular surface.
7. Move along the humerus toward the superior border to examine the articular surface of the supraspinatus muscle.
8. Continue past the supraspinatus muscle to examine the posterior humeral head and bare area.
9. Inspect and probe the humeral head articular surface.
10. Inspect and probe the anterior labrum.
11. Inspect the subscapularis recess and insertion.
12. Inspect the capsular attachment to the humerus (humeral avulsion of the glenohumeral ligament).