



Operative performance outcomes of a simulator-based robotic surgical skills curriculum

William Gerull¹ · Ahmed Zihni¹ · Michael Awad¹

Received: 16 July 2019 / Accepted: 28 October 2019 / Published online: 15 November 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Background Robotic technology has become increasingly prevalent throughout modern surgical practice as surgical training programs are determining how to best include this robotic training in their curricula. In this study, we sought to measure changes in performance and workload metrics in the live operative setting following completion of a novel simulator-based robotic skills curriculum.

Methods 31 surgical residents naïve to robotic platforms were recruited. They first participated in a live robotic case and had a baseline assessment using RO-SCORE, a robotic modification of the O-SCORE tool, and self-assessed their workload using the NASA Task Load Index (NTLX). Subjects then completed the curriculum, created by an expert panel, on a da Vinci Skills Simulator to pre-set proficiency goals. Subjects were encouraged to train on their own time and could complete the curriculum in one or more sittings, in a 1-month time period. Subjects were then assessed after another live case. Data were analyzed using paired Student's *t* test.

Results Completion of the curriculum was associated with significant RO-SCORE improvements in operative performance across all domains including Camera Control (pre-curriculum mean: 1.9; post-curriculum mean: 4.8; $p < 0.001$), Needle Control (pre-curriculum mean: 1.7; post-curriculum mean: 4.4; $p < 0.001$), Tissue Handling (pre-curriculum mean: 2.0; post-curriculum mean: 4.4; $p < 0.001$). There were significant reductions in all NTLX workload domains including Physical Demand (pre-curriculum mean: 5.2; post-curriculum mean: 2.1; $p < 0.001$), and Frustration (pre-curriculum mean: 6.4; post-curriculum mean: 1.42; $p < 0.001$).

Conclusions We describe a feasible robotic simulation curriculum that requires a reasonable amount of time and is self-directed. Significant improvements were seen across all performance metrics and subjective operator workload. Importantly, this is translated to the clinical environment. These types of curricula will be necessary for improving the skills and confidence of trainees and attending surgeons as robotic technology becomes more pervasive.

Keywords Robotic · Curriculum · NTLX · RO-SCORE · O-SCORE · Surgery

Robotic technology has become pervasive throughout modern surgical practice. Since the introduction of the da Vinci Surgical System (dVSS) (Intuitive Surgical, Incorporated; Sunnyvale, CA) to general surgery in 2000, robotic surgery has been rapidly incorporated into many specialties such as gynecology, urology, colorectal, and bariatrics [1]. As a result, surgical training programs across multiple disciplines must adapt to include robotic platform training in their

curriculum. This training adaptation must occur due to the added complexity of robotic surgery—there is a separation of the surgeon's console and the patient, communication difficulties, 3D instead of 2D imaging, limited demonstration capabilities, and a lack of haptic feedback [2]. These new complexities, in addition to our group's own study, have shown that the transferability of skills between laparoscopic and robotic platforms is limited [3]. Thus, a novel, specialized and dedicated robotic-specific skills curricula is required for surgical trainees. Often, surgical trainees have little access to the complete robotic surgery systems for training purposes. Simulator-based robotic training offers a safe alternative training option that overcomes most logistical and patient-safety hurdles. There are multiple options for

✉ William Gerull
wgerull@wustl.edu

¹ Department of Minimally Invasive Surgery, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8109, St. Louis, MO 63110, USA

simulator-based robotic training such as the da Vinci skills simulator by Intuitive Surgical Inc., dV-Trainer from Mimic Technologies Inc., RoSS by Simulated Surgical Sciences LLC, and Robotix Mentor from Symbionix [4]. But, there is still a need for structured curricula to be developed alongside these simulators that focuses on the transferability of simulated robotic skills to the live operative environment [5].

To describe the impact of education curricula, the Kirkpatrick hierarchy of educational outcomes is routinely used in the academic setting. Kirkpatrick's outcome levels applied to medical education research include (1) learner satisfaction; (2) changes in attitudes, knowledge, and skills; (3) changes in behaviors; (4) changes to the care system or patient outcomes. Medical education research should strive to capture outcomes at the Kirkpatrick Level 4, as this level has the highest likelihood for sustained changes in behaviors and improved patient outcomes [6]. To date, most reports of robotic curricula have reached a Kirkpatrick Level 1 or Level 2 of impact by demonstrating robotic simulation curriculum can improve robotic skills in a dry lab setting [7]. Our study aims to reach a higher Kirkpatrick Level of impact by demonstrating improved robotic performance outcomes in a live operative setting associated with a novel simulator-based robotic skills curricula.

Prior surgical research has developed metrics to accurately assess surgical trainee operative competency. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE) is used to assess operative performance and has been successfully applied to surgical simulation tasks. The National Aeronautics and Space Administration-Task Load Index (NTLX) is a validated tool to assess mental workload that has been utilized in many surgical studies [8–10]. Together, the O-SCORE and NTLX are valid forms of training evaluation that can accurately describe changes in a surgeon's performance and workload. In this study, our group introduced a simulator-based robotic skills curriculum and quantified improvements in performance metrics and subjective workload. We developed our simulator-based robotic skills curriculum by consensus of a cross-specialty, expert panel that evaluated a matrix of all available modules. The panel decided to create a curriculum that had commonalities across specialties, was easily replicable and was feasible for trainees to complete with limited time availability. We hypothesize that the introduction of a simulator-based robotic skills curriculum is associated with significant improvements in workload and performance metrics in the live robotic operative setting.

Materials and methods

Subject recruitment

Under an IRB-approved protocol, 31 subjects were recruited and informed consent was obtained in this prospective

cohort pilot study. The subjects were surgical residents in general surgery, urology, and obstetrics and gynecology novice to robotic platforms.

Performance metrics

The attending surgeon assessed subject performance by using a robot-specific modification of the O-SCORE after each case [11, 12]. Our institution modified the original O-SCORE evaluation to better reflect robot-specific surgical characteristics (Fig. 1). A Likert scale scoring system was used, from 1 to 5, using anchor questions of “Complete hands on guidance” to “I did not need to be there” in regard to activities related to each domain. The subject's workload was assessed following each case using the NTLX survey, a tool that allows subjects to self-assess their performances on six sub-scales including Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration [13–15].

Pre-training

All subjects participated in a live robot-assisted laparoscopic surgical (RALS) case prior to reaching proficiency on the novel da Vinci Skills Simulator curriculum. Immediately after the case, the attending surgeon completed the RO-SCORE and the subject completed the NTLX.

Robotic simulator training

After completion of the RALS case, all subjects trained to pre-set proficiency goals on a da Vinci Skills Simulator with a novel skills curriculum. These tasks included: Camera Targeting—Level 2, Energy Dissection—Level 1, Energy Switching—Level 2, Ring and Rail—Level 2, Ring Walk—Level 3, Suture Sponge—Level 3, Thread the Rings, and Tubes. This novel skills curriculum was selected from on-board tasks included with the dVSS software. An expert panel decided the tasks and the difficulty level of each task that was required. The scoring goals that determined proficiency for each required task were based on built-in dVSS software metrics. The dVSS was available to subjects for practice during normal business hours in the Surgical Skills Laboratory at the Washington University Institute for Surgical Education (WISE). Each task listed takes approximately 5 min to complete, with approximately four attempts necessary to reach proficiency. This training took, on average, 3 h to complete. The subjects were given 1 month to complete all tasks to proficiency before moving onto the post-training case.

The Modified Ottawa Surgical Competency Operating Room Evaluation for Surgical Robotic Cases

Trainee:

Procedure:

Date:

Scale

1—"I had to do"—*i.e.*, Requires complete hands on guidance, did not do, or was not given the opportunity to do

2—"I had to talk them through"—*i.e.*, Able to perform tasks but requires constant direction

3—"I had to prompt them from time to time"—*i.e.*, Demonstrates some independence, but requires intermittent direction

4—"I needed to be in the room just in case"—*i.e.*, Independence but unaware of risks and still requires supervision for safe practice

5—"I did not need to be there"—*i.e.*, Complete independence, understands risks and performs safely, practice ready

1. Overall Technical performance	1	2	3	4	5	
Efficiently performs steps, avoiding pitfalls and respecting soft tissues						
a) Camera control	1	2	3	4	5	N/A
b) Energy control	1	2	3	4	5	N/A
c) Needle control/suturing	1	2	3	4	5	N/A
d) Tissue handling/respect for tissue	1	2	3	4	5	N/A
e) Instrument control/wrist movement	1	2	3	4	5	N/A
2. Efficiency and flow	1	2	3	4	5	
Obvious planned course of procedure with economy of movement and flow						
3. Communication	1	2	3	4	5	
Professional and effective communication/utilization of staff						
Other comments/feedback:						
Faculty Signature:						

Fig. 1 Robotic-specific modification of the Ottawa surgical competency operating room evaluation

Post-training

Upon reaching proficiency on all tasks of the novel dVSS curriculum, subjects participated in a live RALS case

with the same attending surgeon that was present for the pre-training RALS case. Immediately, after the case, the attending surgeon completed the RO-SCORE and the subject completed the NTLX.

Data analysis

RO-SCORE and NTLX scores from the pre-training and post-training live RALS cases were compared with paired Student's *t* test.

Results

Thirty-one subjects completed the study, which consisted of a pre-training case, robotic curriculum and post-training case. Completion of the curriculum required an average of 2.3 h. There was a statistically significant improvement seen in all RO-SCORE domains (Fig. 2): Camera Control (Pre-curriculum mean: 1.93, STD: ± 0.78 ; Post-curriculum mean: 4.80, STD: ± 0.42 ; $p < 0.001$), Energy Control (Pre: 1.41 ± 0.50 ; Post: 4.48 ± 0.57 ; $p < 0.001$), Needle Control (Pre: 1.73 ± 0.74 ; Post: 4.38 ± 0.66 ; $p < 0.001$), Tissue Handling (Pre: 1.97 ± 0.76 ; Post: 4.44 ± 0.56 ; $p < 0.001$), Instrument Control (Pre: 1.94 ± 0.65 ; Post: 4.56 ± 0.62 ; $p < 0.001$), Visuospatial (Pre: 1.87 ± 0.68 ; Post:

4.03 ± 0.73 ; $p < 0.001$), Efficiency (Pre: 2.07 ± 0.78 ; Post: 4.52 ± 0.59 ; $p < 0.001$), Communication (Pre: 1.91 ± 0.57 ; Post: 4.29 ± 0.68 ; $p < 0.001$) and Overall (Pre: 2.06 ± 0.85 ; Post: 4.35 ± 0.69 ; $p < 0.001$).

There was a statistically significant reduction seen in all NTLX domains (Fig. 3). This correlated to a decreased workload in all domains; a higher score meant a higher demand in the domain, a lower score meant a lower demand in the domain. For the performance domain, a higher score meant closer to failure, a lower score meant closer to perfect. The significant reductions seen in all NTLX workload ratings were: Mental Demand (Pre-curriculum mean: 6.70, STD: ± 1.5 ; Post-curriculum mean: 3.08, STD: ± 1.52 ; $p < 0.001$), Physical Demand (Pre: 5.23, STD: ± 2.04 ; Post: 2.09, STD: ± 0.99 ; $p < 0.001$), Temporal Demand (Pre: 4.06, STD: ± 1.49 ; Post: 2.05, STD: ± 0.9 ; $p < 0.001$), Performance (Pre: 6.00, STD: ± 1.65 , Post: 2.94, STD: ± 1.59 ; $p < 0.001$), Effort (Pre: 6.83, STD: ± 1.45 ; Post: 3.34, STD: 1.75; $p < 0.001$), Frustration (Pre: 6.44, STD: ± 1.39 ; Post: 1.42, STD: ± 1.01 ; $p < 0.001$).

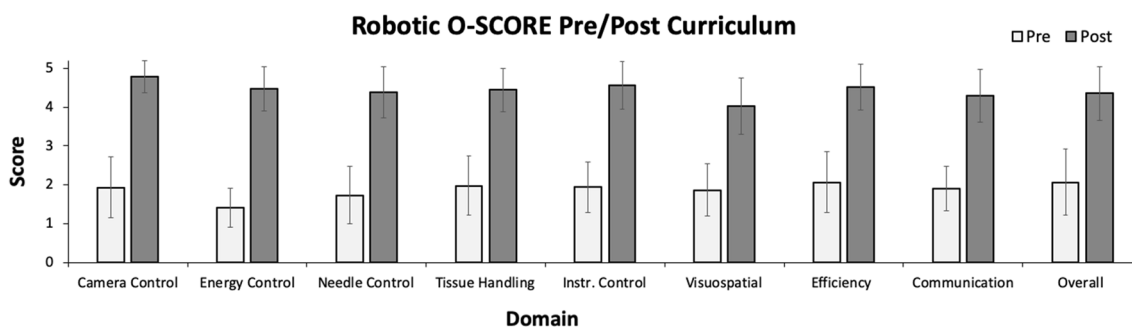
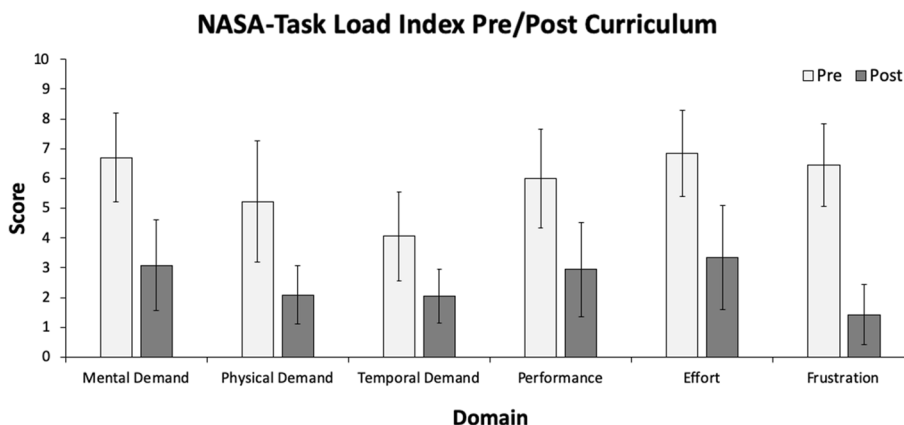


Fig. 2 Mean RO-SCORE domain scores of operations performed by trainees before and after completing the curriculum

Fig. 3 Mean NTLX domain scores of operations performed by trainees before and after completing the curriculum



Discussion

In this study, we have shown that a simulator-based robotic skills curriculum is associated with significant improvements in performance and workload in a live operative setting. While previous literature has shown the utility of robotics curricula in a simulated setting, we have demonstrated a feasible curriculum that translates skills acquisition to actual clinical performance improvement.

The improvements we have demonstrated allow for a trainee to focus on higher-order learning objectives in the operating room such as operative steps, complication anticipation and management, challenging anatomy, etc. because the basic skills of operating the robotic technology is achieved before a trainee steps in the room. It should also allow for safer and more efficient progression through a procedure. Finally, we have found anecdotally that supervising faculty will engage a trainee who has completed the curriculum in a greater portion of a robotic procedure than they might otherwise have, thus enhancing learner satisfaction and engagement.

Having a curriculum that is feasible and replicable is essential for its adoption. As the curriculum was developed by a multi-disciplinary expert panel, it has the benefit of being generalizable across a wide variety of specialty training programs. The curriculum is self-guided, allowing learners to complete it in one sitting or multiple sittings which can alleviate the administrative burden on training program directors and their staff. Because it uses built-in modules provided by the manufacturer, the curriculum is replicable, allows for individuals to track progress, and metrics are provided as objective target goals. In addition, the curriculum can be completed in as little as 2–3 h, which our trainees felt was feasible within the confines of their other clinical obligations.

The findings that we have shown represent one of the first assessments of live operative skills after a training intervention using a robotic simulator. These metrics are important in demonstrating the utility of simulation-based robotic surgical training. Our outcomes fit into the description of the Kirkpatrick Level 3 for educational outcomes, changing the clinical practice of learners through robotic curricula. Capturing patient-centered metrics and live operative performance serve as the highest demonstrable level of impact. Our study further increases the validity of simulator-based robotic training by reporting on subjective operative skill improvements.

A potential limitation of this study is the variability presented by assessment in a live clinical setting. The nature and types of the procedures varied, as did their complexity, and the level of involvement by the trainee. Nonetheless, the positive effect of the curriculum was seen across

the board despite this variability. For future studies, we may video record the performance so that raters could be blinded to the status of the learner, pre-training or post-training, and could include multiple independent raters.

In summary, we have provided evidence to support that training robotic surgeons on simulator-based robotic skills curriculum is associated with significant improvements in workload and performance metrics in the live robotic operative setting. As surgical training programs continue to adopt robotic technologies, our novel robotic surgery skills curriculum can be incorporated into training programs to improve live robotic surgical skills. These findings suggest that the growing field of robotic surgical training should focus on refining simulator-based surgical curricula while using metrics that capture live robotic performance. This study may also help guide the design of future evaluations of simulator-based curricula that incorporate live operative performance.

Funding None.

Compliance with ethical standards

Disclosures Michael Awad: Intuitive research grant (on unrelated project); speaker's bureau. William Gerull and Ahmed Zihni have no conflicts of interest or financial ties to disclose.

References

1. Brody F, Richards NG (2014) Review of robotic versus conventional laparoscopic surgery. *Surg Endosc* 28:1413–1424
2. Amodeo A, Linares Quevedo A, Joseph JV, Belgrano E, Patel HRH (2009) Robotic laparoscopic surgery: cost and training. *Minerva Urol Nefrol* 61:121–128
3. Zihni A, Ge T, Ray S, Wang R, Liang Z, Cavallo JA, Awad MM (2016) Transfer and priming of surgical skills across minimally invasive surgical platforms. *J Surg Res* 206:48–52. <https://doi.org/10.1016/j.jss.2016.06.026>
4. Tanaka A, Graddy C, Simpson K, Perez M, Truong M, Smith R (2016) Robotic surgery simulation validity and usability comparative analysis. *Surg Endosc* 30:3720–3729. <https://doi.org/10.1007/s00464-015-4667-y>
5. Mills JT, Hougren HY, Bitner D, Krupski TL, Schenkman NS (2017) Does robotic surgical simulator performance correlate with surgical skill? *J Surg Educ* 74:1052–1056. <https://doi.org/10.1016/j.jsurg.2017.05.011>
6. Sullivan GM (2011) Deconstructing quality in education research. *J Grad Med Educ* 3:121–124. <https://doi.org/10.4300/JGME-D-11-00083.1>
7. Yardley S, Dornan T (2012) Kirkpatrick's levels and education "evidence". *Med Educ* 46:97–106
8. Guru KA, Shafiei SB, Khan A, Hussein AA, Sharif M, Esfahani ET (2015) Understanding cognitive performance during robot-assisted surgery. *Urology* 86:751–757. <https://doi.org/10.1016/j.urology.2015.07.028>
9. Ruiz-Rabelo JF, Navarro-Rodriguez E, Di-Stasi LL, Diaz-Jimenez N, Cabrera-Bermon J, Diaz-Iglesias C, Gomez-Alvarez M,

- Briceño-Delgado J (2015) Validation of the NASA-TLX Score in ongoing assessment of mental workload during a laparoscopic learning curve in bariatric surgery. *Obes Surg* 25:2451–2456. <https://doi.org/10.1007/s11695-015-1922-1>
10. Yurko YY, Scerbo MW, Prabhu AS, Acker CE, Stefanidis D (2010) Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simul Healthc* 5:267–271. <https://doi.org/10.1097/SIH.0b013e3181e3f329>
 11. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ (2012) The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. *Acad Med* 87:1401–1407. <https://doi.org/10.1097/ACM.0b013e3182677805>
 12. MacEwan MJ, Dudek NL, Wood TJ, Gofton WT (2016) Continued validation of the O-SCORE (Ottawa Surgical Competency Operating Room Evaluation): use in the simulated environment. *Teach Learn Med* 28:72–79. <https://doi.org/10.1080/10401334.2015.1107483>
 13. Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv Psychol* 52:139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
 14. Stefanidis D, Wang F, Korndorffer JR, Dunne JB, Scott DJ (2010) Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload. *Surg Endosc* 24:377–382. <https://doi.org/10.1007/s00464-009-0578->
 15. Dulan G, Rege RV, Hogg DC, Gilberg-Fisher KM, Arain NA, Tesfay ST, Scott DJ (2012) Proficiency-based training for robotic surgery: construct validity, workload, and expert levels for nine inanimate exercises. *Surg Endosc* 26:1516–1521. <https://doi.org/10.1007/s00464-011-2102-6>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.