

Resident self-assessment versus faculty assessment of laparoscopic technical skills using a global rating scale

Sarah E. Peyre¹, Heather MacDonald², Laila Al-Marayati², Claire Templeman², Laila I. Muderspach²

¹ Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, USA

² Department of Obstetrics and Gynecology, Los Angeles County Medical Center, University of Southern California, USA

Correspondence: Sarah E. Peyre, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, USA
Email: speyre@partners.org

Accepted: May 17, 2010

Abstract

Objectives: Accurate self assessment of laparoscopic technical skill development can accelerate a trainee's learning growth and guide life-long development for surgeons. The aim of this study was to measure the accuracy of obstetrical and gynecology resident self assessments as compared to faculty assessments of laparoscopic technical skill performance using a technical skill global rating scale evaluation.

Methods: 37 residents completed four laparoscopic procedures as part of a porcine simulation teaching laboratory: tubal ligation, ovarian cystectomy, salpingostomy and hysterectomy. At the conclusion of the session, residents and faculty completed global rating scale assessments for all laparoscopic procedures completed. Mann-Whitney U analysis was used to compare faculty and resident results.

Results: Overall residents self assessment was lower than

faculty evaluation (148 cases). Post Graduate Year 1's (n=8/n=31 cases) reported significantly lower scores on 5 of 10 items. Post Graduate Year 2's (n=8/n=35 cases) reported significantly higher scores in 6 of 10 items. Post Graduate Year 3's (n=10/n=39 cases) reported lower scores on all ten items (p<0.05). Post Graduate Year 4's (n=11/n=43 cases) reported significantly lower scores in 9 of 10 items.

Conclusions: Residents are more critical of their performance than faculty when using a global rating scale during a laparoscopic simulation teaching session. Implications of inaccurate self perceptions can slow learning development of trainees. Further research is needed to examine the accuracy of resident self perceptions and faculty assessments of performance in the operating room.

Keywords: Global rating scale, self-assessment, simulation, technical skill

Introduction

Laparoscopic skill competency is a central requirement of gynecological residency training.¹ The development of procedural knowledge and the ability to perform the technical tasks that are associated with each laparoscopic procedure requires modeling, deliberate practice and feedback.^{2,3} For adult learners, the ability to accurately self assess performance can facilitate additional learning and growth.^{4,5} One of the aims of training programs is to nurture the skills of residents to be life-long learners.⁶ Self reflection and valid assessment of performance is necessary for this process.

Educational learning theory describes the importance of accurate self assessment.⁷ Stemming from this assertion,

there have been several studies in the medical education literature that have reported on the impact of residents' inability to accurately assess their performance.^{6,8,9} Early research focused on cognitive knowledge rather than technical skill,^{6,9,10} but in the last five years there have been increasing reports of resident technical skill self-assessment in the operating room and in simulated environments.¹¹⁻¹⁷ Findings from these studies have focused on improving resident self-assessment, and the reliability and validity of instruments in measuring true and perceived performance.

To fully understand this issue we need to also have definitive "gold standard" assessment mechanism. Currently, this standard is assumed by the evaluations of faculty, but is

noted by many to have limitations.¹⁵

The American Board of Obstetrics and Gynecology and the Council on Resident Education in Obstetrics and Gynecology have initiated comprehensive in-service and board examinations that measure cognitive knowledge and decision making; but measurement of technical skill relies solely on the individual training programs to certify competency. There have been several attempts by obstetrics and gynecology faculty to develop objective assessment instruments that are reliable and valid. One of these instruments is the Objective Structured Assessment of Technical Skills (OSATS) global rating scale, originally developed for use in general surgery¹⁸, and adopted by several obstetrics and gynecology training programs as an instrument of objective assessment.^{19,20} This 10-item Likert scale global rating instrument measures aspects of technical skill, anatomy and preparation. This instrument has strengths in application because of its minimal rater training and feasibility for faculty evaluations.

This study investigated the following questions: (1) Do residents accurately self assess laparoscopic technical skill as compared to faculty evaluators using a global rating scale instrument? (2) Do residents at different stages in their training and development have differing accuracy in their self assessments of their laparoscopic technical skill?

Methods

Participants

Thirty-seven (post graduate years 1 to 4) obstetrics and gynecology residents from the University of Southern California and Los Angeles County Medical Center participated in this study. Four of the participants were male and 33 were female. This study was a retrospective review of collected evaluation data from five laparoscopic simulation teaching sessions. There were no participants who were eligible that declined to participate. This research protocol received ethical approval by the Institutional Review Board at the University of Southern California.

Task

Prior to the beginning of a laparoscopic porcine laboratory, residents participated in a 30 minute didactic lecture that focused on teaching the planning for surgery, instrumentation, dealing with complications, and general technical considerations of laparoscopic surgery. Residents from the same training level were assigned in pairs to an animal. A faculty preceptor was assigned to each group to provide guidance and feedback on performance during the laboratory. All groups were allowed two hours to practice skills and asked to complete the procedures of tubal ligation, ovarian cystectomy, salpingostomy and hysterectomy. Residents were asked to share technical responsibilities for each case which enabled each of the residents to engage and complete a significant amount of each procedure. Residents were encouraged to ask questions and clarification from

their faculty preceptors on their technical skill and the elements of the procedures.

Assessment

At the conclusion of each laboratory session, each resident was asked to complete a self assessment of their laparoscopic technical skill for each of the four procedures. The instrument was a 10 item global rating instrument that used a five point Likert scale with descriptive anchors. The 10 items included preparedness, respect for tissue, time and motion, knot tying, instrument handling, knowledge of instruments and suture, use of assistants, anatomy and knowledge of specific procedure. This assessment form also asked for an overall performance rating ranging from substandard performance to the level of competent to perform without supervision. Faculty physicians were also asked to fill out this same instrument for each of the two residents that they observed for all four procedures. The same faculty were used during the five teaching sessions and they were all previously trained in using the instrument to evaluate resident performance. Average resident and faculty scores on the 5-point Likert scale, by training class, were used for Mann-Whitney U analysis to measure significant differences.

Results

Faculty and resident assessment data was analyzed for 148 procedures completed during five porcine laparoscopic technical skill sessions. Results showed that overall residents self assessment of technical skills were lower than faculty evaluations with two items being significant; preparedness ($p<0.01$) and knot tying ($p<0.01$). Results were grouped by post graduate year (PGY) class and the technical skill of each procedure to analyze global performance. PGY 1 residents ($n=8$ residents/ $n=31$ completed case evaluations) reported significantly lower scores on 5 of 10 categories: preparedness ($p=0.01$), respect for tissue ($p=0.02$), instrument handling ($p=0.01$), knowledge of instruments/suture ($p=0.02$) and overall performance ($p=0.04$). PGY 2 residents ($n=8$ residents/ $n=35$ completed case evaluations) reported significantly lower scores in 6 of 10 categories: respect for tissue ($p<0.01$), time and motion ($p<0.01$), knot tying ($p=0.03$), instrument handling ($p=0.01$), use of assistants and anatomy ($p=0.03$). PGY 3 residents ($n=10$ residents/ $n=39$ completed case evaluations) reported significantly lower score on all ten items of the scale ($p<0.03$) (Figure 1). PGY 4 residents ($n=11$ residents/ $n=43$ completed case evaluations) reported significantly lower scores in 9 of 10 categories: respect for tissue ($p=0.02$), time and motion ($p<0.01$), knot tying ($p<0.01$), instrument handling ($p<0.01$), knowledge of instrument ($p<0.01$), use of assistants ($p<0.01$), anatomy ($p=0.02$), knowledge of specific procedure ($p<0.01$) and overall performance ($p=0.02$). All results are displayed in Table 1. Overall the obstetrics and gynecology residents who participated in this study self assessed their own technical skills to be lower than faculty preceptor assessments. By resident training class, differences between faculty and residents were significant. These results also

indicate that the more senior the resident is the less likely they will report an accurate assessment of their technical skill. In other words, the gap between self assessment and

faculty assessment of laparoscopic surgical skill widens when looking at senior residents as compared to junior residents.

Table 1. Faculty and resident technical skill assessment results using a global rating scale

Technical Skills	PGY 1 (21)	Faculty (20)	p-value	PGY 2 (33)	Faculty (29)	p-value	PGY 3 (38)	Faculty (32)	p-value	PGY 4 (42)	Faculty (30)	p-value
Preparedness	3.47	4.44	0.01	4.07	3.50	p>0.05	3.82	4.83	p<0.01	4.79	5.00	p>0.05
Respect for tissue	3.14	3.80	0.01	3.82	3.35	p>0.01	3.68	4.34	p<0.01	4.45	4.67	0.02
Time and motion	3.00	3.45	p>0.05	3.79	3.14	p<0.01	3.58	4.18	p<0.01	4.24	4.50	p<0.01
Knot tying	2.20	3.25	p>0.05	3.66	4.00	0.04	3.21	3.80	p<0.01	3.57	4.52	p<0.01
Instrument handling	2.86	3.35	0.01	3.79	3.31	0.01	3.58	4.12	p<0.01	4.17	4.60	p<0.01
Knowledge of instruments	2.81	3.35	0.02	3.76	3.62	p>0.05	3.84	4.24	0.02	4.33	4.60	0.01
Use of assistants	3.10	3.40	p>0.05	3.78	3.35	0.02	3.68	4.21	p<0.01	4.29	4.47	0.02
Anatomy	3.29	3.60	p>0.05	3.79	3.40	0.03	3.66	4.30	p<0.01	4.38	4.63	0.02
Knowledge of procedure	2.86	3.30	p>0.05	3.76	3.40	p>0.05	3.68	4.25	p<0.01	4.41	4.67	0.01
Overall performance	2.89	3.25	0.04	2.89	3.25	p>0.05	3.63	4.28	p<0.01	4.29	4.53	0.02

Discussion

The development of expertise for any skill requires the ability of a practitioner to recognize their capabilities and limitations.²¹ During residency training, individuals receive performance input for their development and growth. However, if a resident does not receive or accurately process faculty formative or summative evaluations then they can develop an inaccurate perception of their abilities. If a resident has a low self efficacy concerning technical performance, their motivation may decline and their learning development may be negatively affected. Considering the technological advances of the specialty, it is crucial for obstetrics and gynecology residents to accurately self assess laparoscopic performance so they can develop competency and confidence in handling the breadth of procedures required. In addition, faculty need to be accurate in their assessments of resident performance to facilitate growth. The results presented in this study imply that either residents have an inaccurate self perception and/or faculty are not critical of senior residents, therefore not providing accurate assessment.

The medical education literature has examined this issue from various points of view that has implications for this body of research. Anthony found in a study of medical students that those who had the least proficiency tended to be more inaccurate in self-assessment and overestimate their abilities.¹⁰ Results presented here show the opposite results in that the more senior the residents are in their training, the wider the gap is in their self assessment. To improve self assessment for all learners it has been shown that review of videotaped performance and repetition on a simulator are both effective.^{14,17} Mandel and colleagues reported obstetrics and gynecology residents using global rating scales can self-assess their surgical skills with good reliability and validity on bench models.¹⁵

Interestingly, that body of research also showed that residents rate their open skills higher than laparoscopic skills; but all skills were rated lower than faculty observations. Those results parallel those demonstrated in this research. Further investigation is needed on why the task domain of laparoscopic surgery has different technical demands that create different resident self perceptions compared to open procedures.

A limitation of this study is that performance evaluation data was collected during a simulation training session where there were multiple residents performing a single operation. This constraint was largely due to funding for animals and due to number of available faculty to participate. Other limitations include that only a single faculty rater was used for each performance, and performance data was collected using a cross sectional methodology.

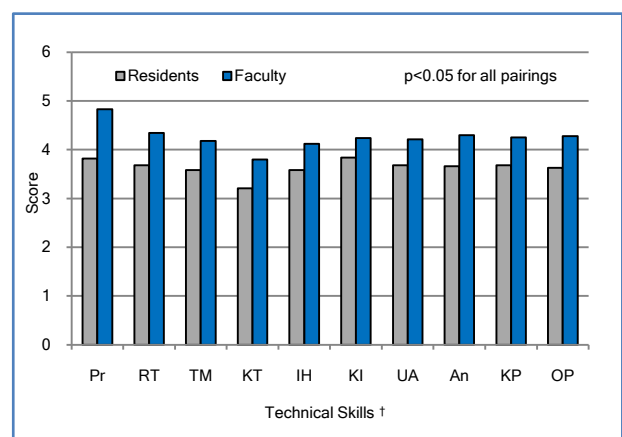


Figure 1. PGY 3 Resident self assessment versus faculty assessment of technical skill performance

[†]Pr, Preparedness; RT, Respect for Tissue; TM, Time and Motion; KT, Knot Tying; IH, Instrument Handling; KI, Knowledge of Instruments; UA, Use of Assistants; An, Anatomy; KP, Knowledge of Procedure; OP, Overall Performance.

A better description of the participants self assessment, and namely development of technical skill would be stronger through a longitudinal examination.

Despite the limitations specific to this body of research, simulation has been widely shown to improve the acquisition of technical skill development in surgical trainees so it is in our opinion that it is an appropriate venue to measure technical skill development. Specific to obstetrics and gynecology, Banks *et al* have shown that compared with apprenticeship teaching alone, a surgical simulator laboratory on laparoscopic tubal ligation improved resident knowledge and performance in the operating room.²²

In a different study, Banks also showed that surgical skills laboratory improved residents' knowledge and performance in the clinical setting. Improvement was greatest for PGY-1 residents.²³ There have also been several studies surrounding the efficacy of simulation training for shoulder dystocia skill development and retention.^{24,25} General surgery simulation research has found similar results.²⁶⁻²⁹

In examining resident performance, if we consider that faculty evaluations are a measure of true performance, then lower resident self perception is a problem that needs to be altered. If residents can not accurately assess their skill level then it will delay their development and possibly prevent them from acquiring new skills. This study occurred looking at resident performance in a simulated environment, but if this perception is consistent then ultimately there is a need to measure if this self perception affects operating room performance. Possibly the role of resident self reflection is to compare resident perceptions with faculty perceptions so that gaps can be identified and residents can develop a more accurate analysis of their surgical skills.

References

1. Gardner R, Raemer DB. Simulation in obstetrics and gynecology. *Obstet Gynecol Clin North Am.* 2008;35:97-127.
2. Norman G, Eva K, Brooks L, Hamstra S. Expertise in medicine and surgery. In: Ericsson K, Charness N, Feltovich P, Hoffman R, editors. *The Cambridge Handbook of Expertise and Expert Performance.* New York: Cambridge University Press; 2006. p. 339-54.
3. Ericsson K. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med.* 2004;79:S70-S81.
4. Spencer J, Jordan R. Learner centred approaches in medical education. *BMJ.* 1999;318:1280-8.
5. Ericsson K. The influence of experience and deliberate practice on the development of superior expert performance. In: Ericsson K, Charness N, Feltovich P, Hoffman R, editors. *The Cambridge handbook of expertise and expert performance.* New York: Cambridge University Press; 2006. p. 683-703.
6. Gordon M. A review of the validity and accuracy of self-assessments in health professions training. *Acad Med.* 1991;12:762-9.
7. Schon DA. *The Reflective practitioner: How professionals think in action.* USA: Basic Books; 1983.
8. Risucci D, Tortolani A, Ward R. Ratings of surgical resident by self, supervisors and peers. *Surg Gynecol Obstet.* 1989;169:519-26.
9. Regehr G, Hoges B, Tiberius R, Lofchy J. Measuring self-assessment skills: an innovative relative ranking model. *Acad Med.* 1996;71:S52-4.
10. Anthony T. A discrepancy in objective and subjective measures of knowledge: do some medical students with learning problems delude themselves? *Med Educ.* 1986;20:17-22.
11. Barnsley L, Lyon P, Ralson S, Hibbert E, Cunningham I, Gordon F, *et al.* Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. *Med Educ.* 2004;38:358-67.
12. Evans SRT, Jackson PG, Czerniach DR, Kalan MMH, Iglesias AR. A stepwise approach to laparoscopic Nissen fundoplication: avoiding technical pitfalls. *Arch Surg.* 2000;135:723-8.
13. Goff B, Nielson P, Lentz G, Chow G, Chalmers R, Fenner D, *et al.* Surgical skills assessment: a blinded examination of obstetrics and gynecology residents. *Am J Obstet Gynecol.* 2002;186:613-7.
14. MacDonald J, Williams RG, Rogers DA. Self-assessment in simulation-based surgical skills training. *Am J Surg.* 2003;185:319-22.
15. Goff B, Mandel L, Lentz G, Vanblaricom A, Oelschlager AM, Lee D, *et al.* Assessment of resident surgical skills: is testing feasible? *Am J Obstet Gynecol.* 2005;192:1331-8; discussion 1338-40.
16. Mandel L, Lentz G, Goff B. Teaching and evaluating surgical skills. *Obstet Gynecol.* 2000;95:783-5.
17. Grober ED, Hamstra SJ, Wanzel KR, Reznick RK, Matsumoto ED, Sidhu RS, *et al.* Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity. *Microsurgery.* 2003;23:317-22.
18. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, *et al.* Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997;84:273-8.
19. Swift S, Carter J. Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty. *Am J Obstet Gynecol.* 2006;195:617-23.
20. Goff B, Lentz G, Lee D, Houmar B, Mandel L. Development of an objective structured assessment of technical skills for obstetric and gynecology residents. *Obstet Gynecol.* 2000;96:146-50.

21. Ertmer PA, Newby TJ. The expert learner: Strategic, self regulated, and reflective. *Instructional Science*. 1996;24:1-24.
22. Banks EH, Chudnoff S, Karmin I, Wang C, Pardanani S. Does a surgical simulator improve resident operative performance of laparoscopic tubal ligation? *Am J Obstet Gynecol*. 2007;197:541.e1-5.
23. Banks E, Pardanani S, King M, Chudnoff S, Damus K, Freda MC. A surgical skills laboratory improves residents' knowledge and performance of episiotomy repair. *Am J Obstet Gynecol*. 2006;195(5):1463-7.
24. Crofts JF, Bartlett C, Ellis D, Hunt LP, Fox R, Draycott TJ. Training for shoulder Dystocia: a trial of simulation using low-fidelity and high-fidelity mannequins. *Obstet Gynecol*. 2006;108:1477-85.
25. Crofts JF, Bartlett C, Ellis D, Hunt LP, Fox R, Draycott TJ. Management of shoulder dystocia: skill retention 6 and 12 months after training. *Obstet Gynecol*. 2007;110:1069-74.
26. Tavakol M, Mohagheghi MA, Dennick R. Assessing the skills of surgical residents using simulation. *J Surg Educ*. 2008;65:77-83.
27. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ*. 2003;327:1032-7.
28. Moorthy K, Munz Y, Jiwanji M, Bann S, Chang A, Darzi A. Validity and reliability of a virtual reality upper gastrointestinal simulator and cross validation using structured assessment of individual performance with video playback. *Surg Endosc*. 2004;18:328-33.
29. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg*. 2004;91:1549-58.