"Reflection-Before-Practice" Improves Self-Assessment and End-Performance in Laparoscopic Surgical Skills Training

Ganni, Sandeep; Botden, SMBI; Schaap, Dennis P.; Verhoeven, Bas H.; Goossens, Richard; Jakimowicz, Jack

DOI
10.1016/j.jsurg.2017.07.030

Publication date
2018

Document Version
Final published version

Published in
Journal of Surgical Education

"Reflection-*Before*-Practice" Improves Self-Assessment and End-Performance in Laparoscopic Surgical Skills Training

Sandeep Ganni, MBBS, MBA,,†,‡ Sanne M.B.I. Botden, MD, PhD,§ Dennis P. Schaap, MD,‡ Bas H. Verhoeven, MD, PhD,§ Richard H.M. Goossens, PhD,* and Jack J. Jakimowicz, MD, PhD*,‡*

*Medisign, Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands; †Department of Surgery, GSL Medical College, Rajahmundry, India; ‡Research and Education, Catharina Hospital, Eindhoven, The Netherlands; and §Department of Paediatric Surgery, Radboud University Medical Center—Amalia Children's Hospital, Nijmegen, The Netherlands

OBJECTIVE: To establish whether a systematized approach to self-assessment in a laparoscopic surgical skills course improves accordance between expert- and self-assessment.

DESIGN: A systematic training course in self-assessment using Competency Assessment Tool was introduced into the normal course of evaluation within a Laparoscopic Surgical Skills training course for the test group ($n = 30$). Differences between these and a control group ($n = 30$) who did not receive the additional training were assessed.

SETTING: Catharina Hospital, Eindhoven, The Netherlands ($n = 27$), and GSL Medical College, Rajahmundry, India ($n = 33$).

PARTICIPANTS: Sixty postgraduate year 2 and 3 surgical residents who attended the 2-day Laparoscopic Surgical Skills grade 1 level 1 curriculum were invited to participate.

RESULTS: The test group ($n = 30$) showed better accordance between expert- and self-assessment (difference of 1.5, standard deviation [SD] = 0.2 versus 3.83, SD = 0.6, p = 0.009) as well as half the number (7 versus 14) of cases of overreporting. Furthermore, the test group also showed higher overall mean performance (mean = 38.1, SD = 0.7 versus mean = 31.8, SD = 1.0, p < 0.001) than the control group ($n = 30$). The systematic approach to self-assessment can be viewed as responsible for this and can be seen as "reflection-*before*-practice" within the framework of reflective practice as defined by Donald Schon.

CONCLUSION: Our results suggest that "reflection-*before*-practice" in implementing self-assessment is an important step in the development of surgical skills, yielding both better understanding of one's strengths and weaknesses and also improving overall performance. (J Surg Ed 75:527-533. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: self-assessment, expert assessment, training, evaluation, laparoscopic cholecystectomy, laparoscopic skills

COMPETENCIES: Patient Care, Medical Knowledge, Practice-Based Learning and Improvement, Systems-Based Practice, Professionalism, Interpersonal and Communication Skills

Correspondence: Inquiries to Sandeep Ganni, Research and Education, Catharina Hospital, Michelangelolaan 2, 5653 EJ, Eindhoven, The Netherlands; e-mail: s.ganni@tudelft.nl

INTRODUCTION

The development of technical skills is crucial for surgical residents and surgeons. Simulation-based training is a very important tool to enhance this competence. Besides supervised teaching by expert, trained surgeons, self-assessment and self-directed learning are key elements in surgical training.[1,2] Several studies have shown that integration of self-assessment is beneficial for the development of a surgeon's career.[3,4]

There is disagreement in terms of the desirable role of self-assessment, between the literature on self-assessment theories and that concerning its real-world implementation in surgical practice. The theoretical literature tends to focus on the use of self-assessment as a means of improving reflective practice and thereby improving the individual's overall professional competence and skills.[5,6] Evaluation of real-world self-assessment in surgical practice often focuses on trying to achieve accordance between expert and self-assessment and a reduction in overestimation of performance.[7]

Although self-assessment has been considered a vital component for professional self-regulation and development for a long time, many studies debate the effectiveness and efficacy of self-assessment in skills training and state that there is room for improvement.[5,8-10] Recently, several authors, such as Ward et al,[11] propose that resolving weaknesses in the methodologies used to evaluate self-assessment would yield a more positive evaluation of self-assessment's efficacy. Because of these improved methodologies, it has been shown that trainees or surgical residents are in fact able to self-assess their weaknesses and strengths similarly to expert assessment.[7,12,13]

Regardless of the field using self-assessment, the ideal is to improve the ability of individual candidates to accurately assess their own ability with the aim to improve their overall performance; to this end, many tools and methodologies have been suggested for the improvement of self-assessment itself.[6,14,15] One of the most important conclusions is that surgical residents assess their own procedural performance more accurately after watching benchmark videos of expert performances and their own performances.[8,16] Stewart et al indicated a concentrated, intense course in procedural skills before evaluation for self-assessment to be more accurate, namely greater accordance between expert- and self-assessment.[17]

This study aimed to determine whether implementing a self-assessment training tool in a validated laparoscopic surgical skills course will improve the accordance between self- and expert assessment.

MATERIALS AND METHODS

Participants

Sixty surgical residents who attended the 2-day Laparoscopic Surgical Skills (LSS) grade 1 level 1 curriculum were invited to participate in 2 centers: Catharina Hospital, Eindhoven, The Netherlands ($n = 27$), and GSL Medical College, Rajahmundry, India ($n = 33$). Their expertise level ranged from postgraduate year 2 to 3. All participants voluntarily enrolled in the study and signed an informed consent before the start of the curriculum. All participants had completed and passed an online examination on the basics of laparoscopic surgery to be eligible for participation in the program. Each participant completed a questionnaire with questions pertaining to demographics, experience in laparoscopic surgery, and time spent preparing for the curriculum.

Assessment tool

The Competency Assessment Tool (CAT) used in this study is an operation-specific assessment tool that was adapted and validated for the laparoscopic cholecystectomy (LC) procedure for use within the LSS curriculum.[18] In this study, it was used as a tool for self- and expert assessment. The CAT evaluation criteria are spread across 3 procedural tasks: (1) exposure of both the cystic artery and cystic duct, (2) cystic pedicle dissection, and (3) resection of the gallbladder from the liver bed. Within these tasks, performance is rated on a five-point, task-specific scale based on the efficient usage of instruments, the handling of tissue with the nondominant hand, errors within each task, and the end-product of each task. A maximum of 48 points can be scored on the CAT assessment, and a total score of 30 or more was considered a pass for the LC course.

Four expert surgeons, 2 from each of the respective locations conducting the curriculum, were invited to participate as expert assessors for both the test and control groups. They all had previous experience in using the CAT form for evaluation. Their laparoscopic surgical experience ranged from 5 to 25 years, each with more than 200 laparoscopic procedures performed as main surgeon. The surgeons were not aware whether the candidates they were assessing had the additional training or not when conducting their assessment.

Protocol

All participants completed the standard training and instructions of the LSS grade 1 level 1 curriculum. During the course, they received an interactive discursive training with experts on the basics of laparoscopic surgery, LC, virtual reality simulators, and box trainers. The participants were divided in 2 groups based on the days they attended courses; into a test group ($n = 30$) and a control group ($n = 30$).

The participants of both groups were instructed by the expert surgeons on the procedural tasks of the LC. Immediately before they performed the procedure, the test group received an additional training session on self-assessment (Fig. 1).

This session totaled 30 minutes in duration and started with the instructor introducing the theoretical meaning and professional benefits observed in the literature of self-assessment. The group was then given the CAT form and instructed to read it. Each criterion was explained in detail by the instructor. The relation between the word-based definitions on the CAT form and their score-based equivalents was explained. The instructor then held a question and answer session to resolve any of the participant's concerns. Where possible, the criteria were accompanied

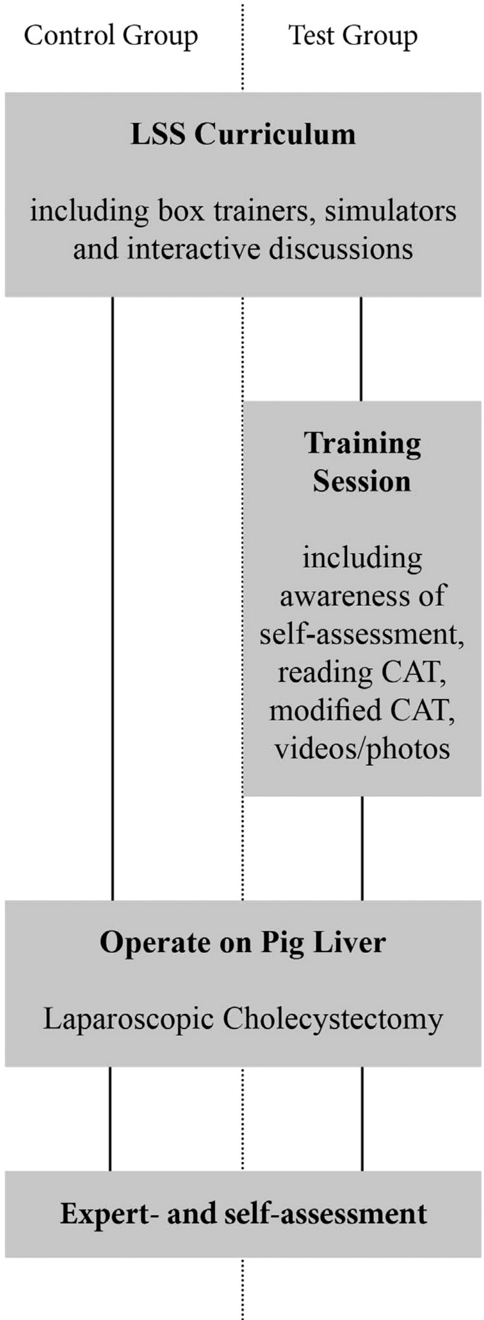

FIGURE 1. An illustration of the training protocol depicting the differences between the test and control groups.

by illustrative videos, showing examples of both good (CAT score of 4) and bad (CAT score of 1) practice, for additional explanation and images of the same were printed overleaf the CAT form for later reference (Fig. 2). As is the current norm, the control group were given the CAT form just before the procedure.

Thereafter, every participant performed the procedural tasks of the LC on a porcine liver placed in a box trainer. The box trainer was placed on a height-adjustable table with an ergonomically correct position of monitors and instruments. The entry ports for the laparoscopic instruments mimicked the incision points in the clinical setting. A fellow participant played the role of surgical assistant during the procedure. The assessors were asked not to provide feedback on the participant's performance during the procedure. Immediately after the procedure, each participant and an expert observer completed a CAT form independently of each other.

Statistical Analysis

The analysis was performed by comparing the differences between the expert- and self-assessment scores between the test and control groups based on the aforementioned criteria within the procedural tasks. Statistical and absolute differences were calculated between expert- and self-assessment scores using MATLAB (R16b), and the obtained data were analyzed and presented using GraphPad Prism (Version 7.00). Because the data were nonparametric, the Mann-Whitney *U* test was used to calculate significant differences between the assessment scores. Other statistical differences were calculated using Graphpad Prism. A $p < 0.05$ was considered statistically significant.

Both the numerical difference and the absolute difference between expert assessment and self-assessment scores were used. When only using one of these measures, clinically relevant correlations or differences could get lost. For example, if part of the trainees score themselves higher and the other half lower than the experts, this is clinically relevant, but the mean of the self-assessment would be equal to the expert assessment. An improvement in the quality of self-assessment can be seen if the numerical difference between self-assessment and expert assessment is closer to zero, when comparing the test group to the control group. If the absolute difference is smaller, this corresponds to an improvement in self-assessment after self-assessment training.

RESULTS

Assessment score

The total overall score given by the experts for the test group in the CAT assessment shows a significantly higher mean than for the control group (mean = 38.1, standard deviation [SD] = 0.7 versus mean = 31.8, SD = 1.0, $p < 0.001$) (Fig. 3). In the control group, 9 participants scored below 30 (regarded as a fail for the course) on the expert assessment, whereas no participant scored less than 30 in the test group.

The same pattern is seen in the self-assessment results, with a mean of 37.6 (SD = 0.6) for the test group and mean of 32.8 (SD = 0.8) for the controls ($p < 0.0001$). On self-assessment, 7 participants of the control group scored less than 30, compared to none in the test group.

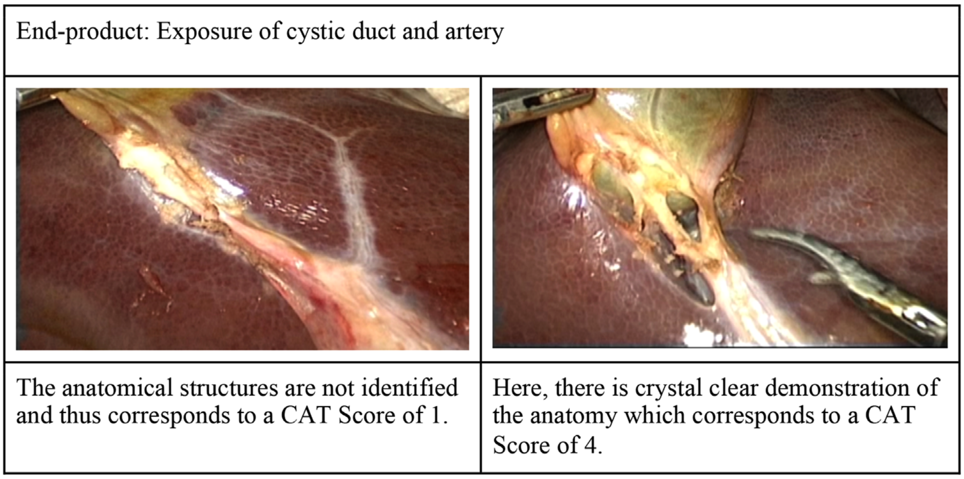

FIGURE 2. An example of the photographic reminders of the CAT scores corresponding to both good and bad practices which were overleaf the modified CAT form given to the test group participants.

The scores of the individual tasks, for both expert- and self-assessment (Table 1), without significant differences in outcome between the test and control group on the separate scored items for both expert- and self-assessment. However, when it comes to the use of tools in the cystic pedicle dissection, the self-assessment was 0.3 points higher on average for the control group compared with the expert assessment, while their expert-assessed performances were the lowest of all (Table 1).

Figure 4 shows the distribution of the differences between the total expert-assessed and the self-assessed CAT scores. The interquartile range and SD between the scores are much smaller in the test group than in the control group. Also, the mean absolute difference between expert- and self-assessment is significantly lower in the test group compared to the control group (1.5 versus 3.83), with a smaller SD (0.2 versus 0.6) ($p = 0.009$) (Table 2). In addition, the number of overestimated performances decreased from 14 in the control group to 7 in the test group (Fig. 4).

Looking at the absolute difference between expert- and self-assessment (Fig. 5) shows that the mean is much lower in the test group -1.5 ± 0.2 versus 3.83 ± 0.6 ($p = 0.009$). Table 2 explains the differences between these values for the grouped items (usage of instruments, tissue handling, errors, and end-product evaluation). Significant differences are seen in all items, except for the usage of instruments, and the items on this subject were scored equal in both groups. As also shown in the table, the calculated difference between self- and expert assessment of tissue handling was 0.0 in the control group; however, the absolute differences shows 1.27.

TABLE 1. Expert-Assessment (EA) and Self-Assessment (SA) Scores on the Separate Scored Aspects: The Mean (standard deviation) Response for Each Criterion on the CAT Form is Shown for Both the Total Group and Test/Control Group Only. Differences Between Expert- and Self-Assessment in Total CAT Scores

	Total EA	Total SA	Test EA	Test SA	Control EA	Control SA
Use of graspers and tools						
Exposure of cystic duct and artery	2.8 (0.58)	2.7 (0.45)	3.0 (0.37)	2.9 (0.25)	2.6 (0.67)	2.6 (0.56)
Cystic pedicle dissection	2.8 (0.69)	2.8 (0.52)	3.2 (0.55)	3.0 (0.56)	2.4 (0.63)	2.7 (0.45)
Resection of gallbladder	3.0 (0.74)	2.8 (0.64)	3.2 (0.66)	3.1 (0.59)	2.8 (0.76)	2.6 (0.56)
Tissue handling						
Exposure of cystic duct and artery	2.9 (0.59)	2.8 (0.56)	3.2 (0.43)	3.1 (0.35)	2.6 (0.61)	2.6 (0.61)
Cystic pedicle dissection	2.9 (0.70)	2.9 (0.58)	3.2 (0.68)	3.0 (0.61)	2.7 (0.64)	2.9 (0.55)
Resection of gallbladder	3.0 (0.65)	3.0 (0.64)	3.3 (0.53)	3.3 (0.48)	2.8 (0.66)	2.6 (0.61)
Errors						
Exposure of cystic duct and artery	2.9 (0.64)	2.9 (0.69)	3.2 (0.50)	3.1 (0.53)	2.7 (0.65)	2.7 (0.75)
Cystic pedicle dissection	2.8 (0.71)	2.9 (0.75)	3.1 (0.63)	3.1 (0.65)	2.5 (0.68)	2.8 (0.81)
Resection of gallbladder	2.6 (1.05)	2.9 (0.86)	2.9 (0.87)	3.1 (0.70)	2.4 (1.17)	2.6 (0.93)
End evaluation						
Exposure of cystic duct and artery	2.9 (0.56)	3.0 (0.58)	3.1 (0.46)	3.2 (0.50)	2.6 (0.55)	2.7 (0.57)
Cystic pedicle dissection	2.9 (0.70)	2.9 (0.54)	3.2 (0.68)	3.0 (0.61)	2.6 (0.61)	2.9 (0.45)
Resection of gallbladder	2.9 (0.80)	3.0 (0.65)	3.2 (0.81)	3.2 (0.25)	2.7 (0.74)	2.8 (0.55)

TABLE 2. Summative Statistics (Mean [Standard Deviation; Standard-Error in the Mean]) for the Numeric and Absolute Difference Between the Expert Assessment (EA) and Self-Assessment (SA) Between the Test and the Control Group. Significant Differences are Calculated Using the Mann-Whitney *U* test, with $p < 0.05$ Considered a Significant Difference

	Difference between EA versus SA		p Value	Absolute difference EA versus SA		p Value
	Test group	Control group		Test group	Control group	
All criteria	0.63 (1.52; 0.28)	−0.97 (4.97; 0.91)	**0.045**	1.5 (1.11; 0.20)	3.83 (3.24; 0.59)	**0.009**
Usage of instruments	0.27 (0.91; 0.17)	−0.10 (1.37; 0.25)	0.207	0.80 (0.48; 0.09)	0.97 (0.96; 0.18)	0.841
Tissue handling and usage of NDH	0.60 (0.50; 0.09)	0.00 (1.74; 0.32)	**0.040**	0.6 (0.50; 0.09)	1.27 (1.17; 0.21)	**0.019**
Errors	−0.20 (0.81; 0.15)	−0.43 (1.94; 0.35)	0.561	0.6 (0.56; 0.10)	1.43 (1.36; 0.25)	**0.011**
End-product evaluation	0.10 (0.80; 0.15)	−0.43 (1.33; 0.24)	**0.047**	0.63 (0.49; 0.09)	1.03 (0.93; 0.17)	0.097

NDH, nondominant hand.
Bold: Statistically significant

DISCUSSION

It is important that doctors and, in particular, surgeons know how they perform during surgical procedures. If a surgeon is not aware of possible hazardous movements or near-incidents, this could result in unnecessary high complication rates. Therefore, it is important that surgeons are assessed on their skills before they perform unsupervised procedures in the clinical setting. When surgeons are accurately aware of their own skills level, with strengths and weaknesses, they know what skills are important to practice more extensively. Therefore, self-assessment could be an important step in the development of surgical skills and enhance patient safety.

This study aimed to assess whether implementing a self-assessment training tool, which includes the latest

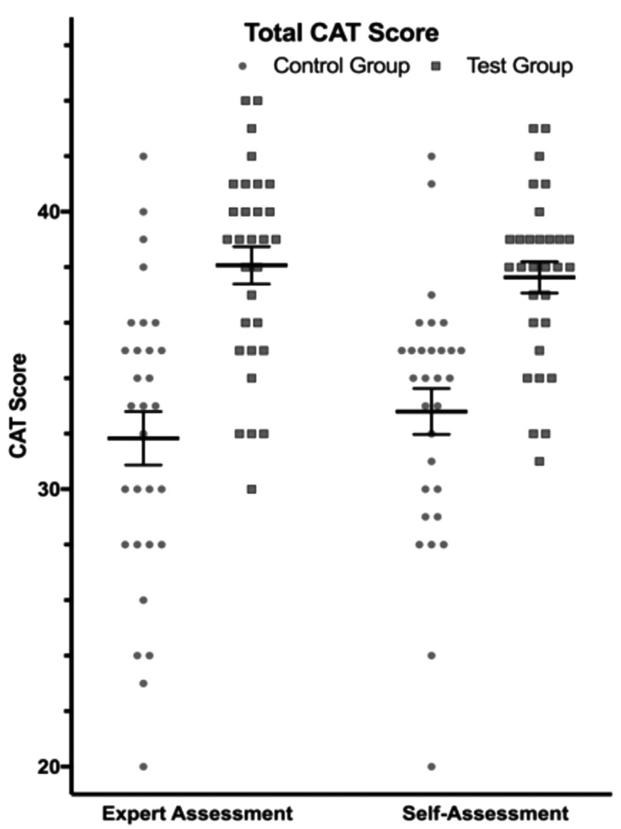

FIGURE 3. The total score on the CAT form for each participant as assessed by an expert.

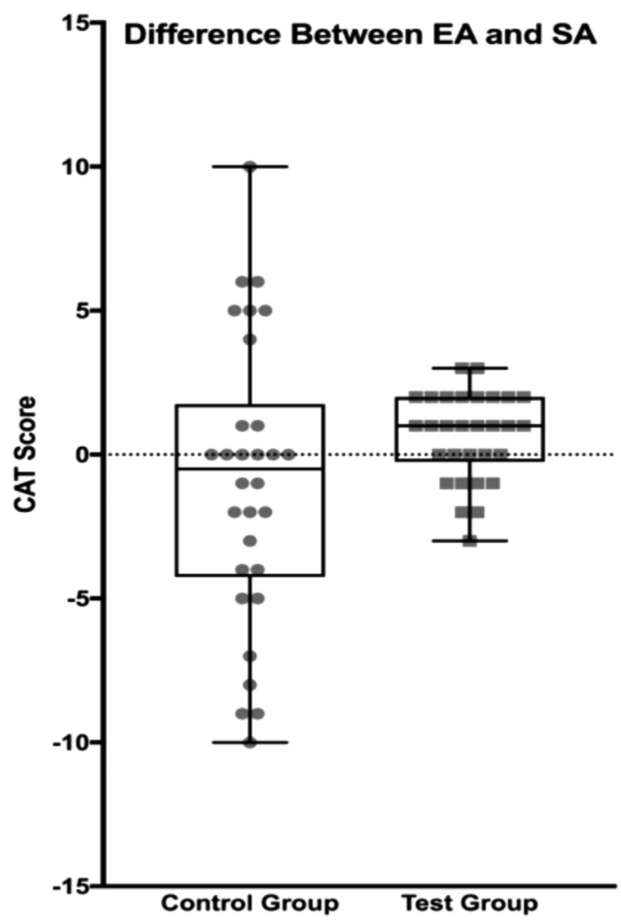

FIGURE 4. Box-plot of the difference between expert- and self-assessment scores, by means of deviation of the individual assessment scores between the 2 groups (p = 0.045).

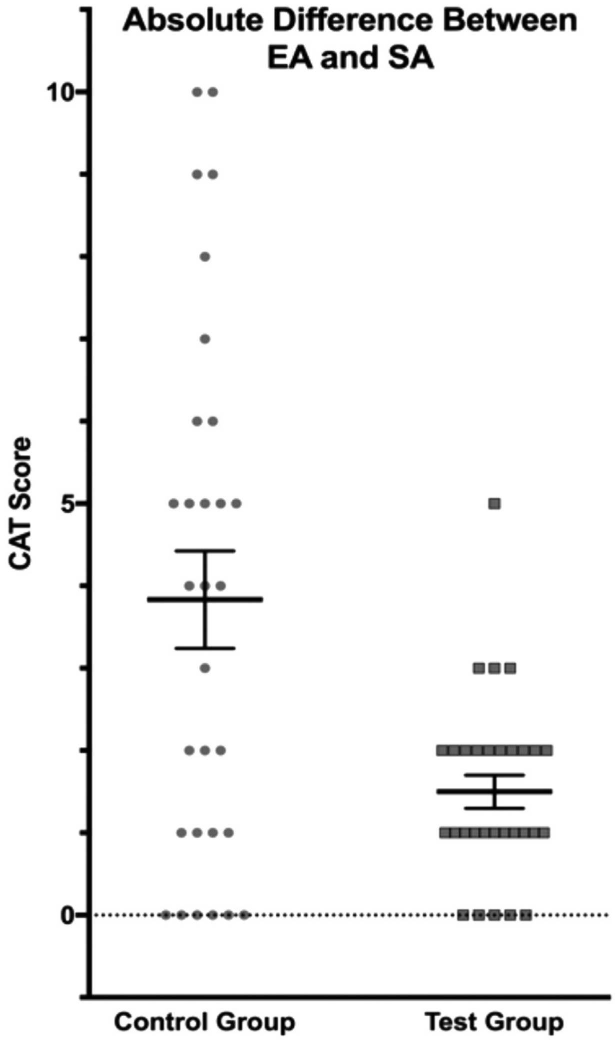

FIGURE 5. The absolute value of the difference between expert- and self-assessment. Individual difference scores are shown as gray points, with the mean ± SEM shown in black.

methodology recommendations, improves the accordance between self- and expert assessment in a validated laparoscopic surgical skills course. This resulted in a single training session immediately before performing the procedure to allow the candidates time to reflect on the assessment form and its criteria based on which their skills were assessed. Additionally, it drew attention to the fact that effective self-assessment was for their long-term professional benefit. Furthermore, the recommendations of previously described literature were implemented through an additional training session by means of videos and photographs of both good and bad practice.[6,12-15]

While it is possible to expect that training in how to use the CAT form would yield the better accordance between expert- and self-assessment seen in the results, it is interesting to note that this systematic approach to implementing self-assessment also improved the candidates' overall performance. Another benefit was a reduction in the number of candidates in the test group who over-reported their performance, suggesting that the training session may have made the participants more aware of their proficiency.

The candidates in the test group may also have benefitted from considering the exact assessment criteria before assessing themselves, resulting in a better understanding of both the criteria themselves and the relationship between their word-based definition and their score-based equivalents. Moreover, the value of both educating and motivating the candidates with the short- and long-term benefits of accurately applying self-assessment to their practice resulted in increased performance outcomes across almost all evaluation criteria.

The concept of reflective practice, as it is currently understood, beholds 3 categories defined by Donald Schon: knowing-in-action, reflection-on-action, and reflection-in-action.[19] However, in implementing this improved approach to self-assessment, it seems we have a form of "reflection-*before*-practice" here. That is, providing the candidates with an understanding of the individual professional benefits of self-assessment, as well as examples of both good and bad practice, during and *before* they self-assess. Furthermore, providing the candidates with a clear objective of the expected outcomes in advance could have created a constant reflection of these objectives during the course of their performance. It is this that appears to have resulted in improved overall performance and increased accordance between expert- and self-assessment.

The scope of this study is limited by its implementation on a relatively basic laparoscopic procedure and assessment. More comprehensive assessment tools such as the observational clinical human reliability assessment (OCHRA) have been highly regarded in assessing not only the competency of skills but also human reliability by means of consequential and nonconsequential errors during a surgical procedure.[20] Further research should be done to investigate whether "reflection-*before*-practice" using assessment tools such as OCHRA in laparoscopic procedures would also improve both competency and human reliability factor of surgeons. Furthermore, establishing whether similar improvements are seen when this approach is applied to a variety of surgical procedures would prove the approach to be an effective way of improving both self-assessment and, importantly, performance.

CONCLUSION

As might have been expected, candidates who receive training in self-assessment in surgical skills training significantly improve their accordance between self- and expert assessment and a reduction in overreporting. Here, however, a second function of the training in self-assessment is seen; the participants improved their overall performance.

Thus, training in self-assessment, seen as "reflection-*before*-practice," can be used to improve not only the accuracy of self-assessment but also the actual surgical performance.

REFERENCES

1. Kostons D, van Gog T, Paas F. Training self-assessment and task-selection skills: a cognitive approach to improving self-regulated learning. *J Learn Instr*. 2012. https://doi.org/10.1016/j.learninstruc.2011.08.004.

2. Kwasnik EM, Carter J. To see ourselves as other see us: self-assessment in surgical residency. *Curr Surg*. 1999. https://doi.org/10.1016/S0149-7944(99)00041-0.

3. Arora S, Miskovic D, Hull L, et al. Self versus expert assessment of technical and non-technical skills in high fidelity simulation. *Am J Surg*. 2011. https://doi.org/10.1016/j.amjsurg.2011.01.024.

4. Gordon MJ. Self-assessment programs and their implication for health professions training. Available at: http://journals.lww.com/academicmedicine/Abstract/1992/10000/Self_assessment_programs_and_their_implications.12.aspx. 1992. Accessed 10.03.17.

5. Boud D. Avoiding the traps: seeking good practice in the use of self assessment and reflection in professional courses. *Soc Work Educ*. 1999. https://doi.org/10.1080/02615479911220131.

6. Boud D. Enhancing Learning Through Self-Assessment. New York: Routledgefarmer; 2005.

7. Ganni S, Chmarra MK, Goossens RHM, et al. Self-assessment in laparoscopic surgical skills training: Is it reliable? *Surg Endosc*. 2016. https://doi.org/10.1007/s00464-016-5246-6.

8. Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Acad Med*. 1999;66:762-769.

9. Arnold L, Willoughby TLEV. Self-evaluation in undergraduate medical education: a longitudinal perspective. *J Med Educ*. 1985;60:21-28.

10. Eva KW, Cunnington JPW, Reiter HI, et al. How can I know what I don't know? Poor self-assessment in a well defined domain *Adv Health Sci Educ*. 2004. https://doi.org/10.1023/B:AHSE.0000038209.65714.d4.

11. Ward M, MacRae H, Schlachta C, et al. Resident self-assessment of operative performance. *Am J Surg*. 2002. https://doi.org/10.1016/S0002-9610(03)00069-02.

12. MacDonald J, Williams RG, Rogers DA. Self-assessment in simulation-based surgical skills training. *Am J Surg*. 2003;185:319-322.

13. Mandel LS, Goff BA, Lentz GM. Self-assessment of resident surgical skills: is it feasible? *Am J Obstet Gynecol*. 2005. https://doi.org/10.1016/j.ajog.2005.07.080.

14. Jamshidi R, LaMasters T, Eisenberg D, et al. Video self-assessment augments development of videoscopic suturing skill. *J Am Coll Surg*. 2009. https://doi.org/10.1016/j.jamcollsurg.2009.07.024.

15. Sarker SK, Hutchinson R, Chang A, et al. Self-appraisal hierarchical task analysis of laparoscopic surgery performed by expert surgeons. *Surg Endosc*. 2006. https://doi.org/10.1007/s00464-005-0312-5.

16. Hu Y, Tiemann D, Brunt LM. Video self-assessment of basic suturing and knot tying skills by novice trainees. *J Surg Educ*. 2013. https://doi.org/10.1016/j.jsurg.2012.10.003.

17. Stewart RA, Hauge LS, Stewart RD, et al. A CRASH course in procedural skills improves medical students' self-assessment of proficiency, confidence, and anxiety. *Am J Surg*. 2007. https://doi.org/10.1016/j.amjsurg.2007.01.019.

18. Miskovic D, Ni M, Wyles SM, et al. Is competency assessment at the specialist level achievable? A study for the national training program in laparoscopic colorectal surgery in England *Ann Surg*. 2013;257:476-482.

19. Schon D. The Reflective Practitioner. How Professionals Think in Action. London: Temple Smith; 1983.

20. Tang B, Hanna GB, Joice P, et al. Identification and categorization of technical errors by observational clinical human reliability assessment (OCHRA) during laparoscopic cholecystectomy. *Arch Surg*. 2004. https://doi.org/10.1001/archsurg.139.11.1215.