The American Journal of Surgery (2012) 204, 724–731

The American Journal of Surgery®

Clinical Science

Are residents accurate in their assessments of their own surgical skills?

Catherine de Blacam, M.R.C.S., M.D.*, Dara A. O'Keeffe, M.R.C.S., Emmeline Nugent, M.R.C.S., Eva Doherty, M.Psych.Sc., Oscar Traynor, M.Ch., F.R.C.S.I.

National Surgical Training Centre, Royal College of Surgeons in Ireland, 123 St. Stephens Green, Dublin, Ireland

KEYWORDS:
Surgical skills;
Self-assessment;
Self-awareness;
Surgical training

Abstract

BACKGROUND: An appropriate level of awareness of one's own technical skill is essential to being an effective self-directed learner. The aim of this study was to analyze surgical residents' self-awareness by examining their ability to predict and assess their own performance in an objective surgical skills examination.

METHODS: Surgical residents' (n = 216) pre-examination self-predicted scores and post-examination self-assessed scores were compared with objectively measured scores. Accuracy of score prediction and self-assessment were compared with resident demographics (age, gender, year of training, and nationality).

RESULTS: Post-examination self-assessed scores correlated with objectively assessed scores (r = .34; $P < .001$). Higher year of training, older age, and non-European nationality were predictive of accuracy in self-prediction and self-assessment.

CONCLUSIONS: Demographic variables that predict more accurate self-awareness of technical skill have been identified. Surgical training programs may use these data to anticipate the trainees who need assistance in developing better self-awareness.

A paradigm shift in surgical education has been created by reductions in the working hours and training time of residents and also by the increasing expectations of governance bodies, the medical profession, and patients. In 2003, the Accreditation Council for Graduate Medical Education mandated restriction of resident working hours in the United States.[1] In Europe, even tighter limitations are in place as a result of the UK Calman reforms and the European Working Time Directive.[2] Although surgical trainees report improved quality of life, concerns about adequate skill acquisition have been raised.[3–6] However, with the overwhelming support of regulatory bodies and the general public, limitation of resident work hours is now firmly established.[7,8]

In response to these changes, surgical training bodies have been obliged to adapt the way in which training is delivered. The result is a curriculum that is structured around proficiency-based performance goals, rather than experiential learning. In this model, technical skills are practiced and built up in a stepwise manner until proficiency criteria are achieved.[9] A significant part of training now takes place outside of the operating room, making use of a variety of platforms such as bench model simulation and

* Corresponding author. Tel.: +353 1 4022701; fax: +353 1 4022459.
E-mail address: catherinedeblacam@rcsi.ie
Manuscript received September 10, 2011; revised manuscript March 4, 2012

0002-9610/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amjsurg.2012.03.003

virtual reality training, as well as cadaver and live animal model surgery. The benefits of such innovations in terms of the trainees' learning curve and ensuring patient safety are indisputable. However, many of these technologies provide little feedback and rely heavily on self-assessment.[10] Thus, the modern surgical training paradigm has placed an increased emphasis on self-directed learning. An appropriate level of awareness of one's own skill is essential to being an effective self-directed learner. Indeed, delivery of surgical care at every level is based on the assumption that surgeons are cognizant of their capabilities and limitations.

All of these factors point to the need for more formal assessment of surgeons' awareness of their technical abilities. As their training advances, surgeons are required to assume the role of lead surgeon, and the ability to confidently and accurately predict one's performance is the key to this transition. With increased emphasis on continued medical education, both trainees and senior surgeons are required to assess their own level of technical skill on an ongoing basis. Thus, self-awareness is an essential personal characteristic for surgical practice and this has been recognized by the Accreditation Council for Graduate Medical Education as a core competency in graduate surgical training.[11,12]

A surgical curriculum that incorporates formal human factors training is well established in the Royal College of Surgeons in Ireland.[13] As residents progress through their basic and specialist training, a series of taught modules and workshops explore the personal characteristics necessary for surgical practice. These include decision making, communication, teamwork, leadership, and self-awareness. In the current study, we sought to examine the level of self-awareness in first- and second-year residents. The study was undertaken in the context of an objectively scored Objective Structured Assessment of Technical Skills (OSATS)-style assessment. We divided technical self-awareness into 2 categories: performance prediction and post-task self-assessment. Before assessment, residents were asked to estimate what score they thought they would achieve and this was used as a measure of their self-prediction ability, indicating their level of confidence in their existing technical skills. To estimate trainees' self-assessment accuracy, they were asked to rate themselves again immediately after having completed the test. Information on a number of demographic factors was collected to investigate whether they were associated with accuracy in technical skill self-awareness.

Methods

Study design and participants

Institutional ethics approval was obtained from the Royal College of Surgeons in Ireland. This was an observational study that was performed in the National Surgical Training Centre at the Royal College of Surgeons in Ireland. All first- and second-year trainees in the National Basic Surgical Training Program were invited to participate in the study. Laboratory-based technical skill assessments were performed as part of their annual Competence Assessment and Performance Appraisal process. Trainees were surveyed before and after the technical skills assessment to evaluate their ability to accurately predict and assess their surgical proficiency. Explanatory information pertaining to the study was supplied to all participants, expressly outlining that their predicted and self-assessed scores would not affect the actual scores they would receive in the assessment. Written informed consent was obtained from all participants.

This study was conducted using a questionnaire that was developed specifically for this purpose, and comprised demographic questions and a self-scoring component. In advance of administration to the whole group, the questionnaire was piloted among a convenience sample of surgical trainees (n = 20). The questionnaires were found to be easily understandable and quick to complete.

Before commencing the formal assessment, trainees were informed of the nature of the tasks on which they were about to be assessed and the potential range of scores achievable. They then were asked to make a numeric prediction of the score they thought they would achieve in each task. After completion of the technical assessments, trainees were asked to estimate what score they thought they had achieved out of the maximum checklist score for each station.

Three separate time-limited stations were created, incorporating skills from all modules that the trainees had completed in the skills laboratory during the preceding year. First-year trainees were asked to complete the following: (1) suture repair of a laceration, (2) excision of a subcutaneous lesion, and (3) incision and closure of a laparotomy. Second-year trainees were asked to complete the following: (1) a bowel anastomosis, (2) ligation of the saphenofemoral junction, and (3) basic laparoscopic skills on a high-fidelity ProMIS simulator (Haptica Ltd., Dublin, Ireland). Other than the laparoscopic simulator, all stations made use of low-fidelity bench models with artificial tissues and vessels. Surgical-grade suture materials and surgical instruments were used in all stations.

Surgical skills assessment

Objective assessment of the trainees' performance was performed during the task by trained attending-level faculty observers, using objective surgical skills assessment (OSSA) checklists. OSSA is a surgical technical skills assessment instrument that is under development in the National Surgical Training Centre and follows the principles of the OSATS,[14] incorporating a task-specific checklist and a global score. This objective instrument was developed specifically for examining technical skills in basic surgical trainees in Ireland. The OSSA checklists were created after a review of each module's learning objectives and were validated by Delphi analysis. The task-specific checklists were examined in a pilot study of 50 trainees for inter-rater reliability, which revealed reliability of .76 to .93, as measured by Cronbach α. Two expert faculty observers on each assessment day were trained in observation and scoring techniques. Also, a random 10% sample of assessment sessions had 2 simultaneous observers to ensure continued adequate inter-rater reliability. Observers scored the trainees as they worked and therefore were not blinded to their identities. However, care was taken to ensure that none of the trainees were known personally to the observers. Each trainee had a different faculty observer for each task and therefore was scored by a total of 3 observers for the assessment.

Statistical analysis

Data were analyzed using the Statistical Package for the Social Sciences (SPSS version 18.0; SPSS, Inc, Chicago, IL). The relationship between objective and self-assessment scores was evaluated using the Pearson correlation coefficient (a linear correlation coefficient, r, such that $-1 \leq r \leq +1$, with stronger positive correlations being closer to $+1$); a *P* value less than .05 was considered significant. These data were represented graphically using scatter plots, with the value of the objective scores determining the position on the x-axis and the self-assessed scores determining the position on the y-axis.

Bland and Altman[15] plots were used to assess the agreement between objective and self-assessment. The Bland and Altman[15] plot, or difference plot, is a graphic method to compare 2 measurement techniques: in this case, objective and self-assessed scores. The x-axis is the mean of the 2 measurements, which is one's best guess as to the correct result, and the y-axis is the difference between the 2 measurements. If the points are scattered broadly above and below zero, this suggests that there is no consistent bias of one approach versus the other. Bland and Altman[15] plots have been used previously in surgical education studies to compare self-assessment and faculty scores.[16,17]

The difference between an individual trainee's self-predicted score (SPS) or self-assessed score (SAS) and objectively assessed scores (OAS) was calculated as follows: SPS-OAS and SAS-OAS. The average SPS-OAS and SAS-OAS was compared with trainee demographics (age, sex, year of training, and nationality) using the Pearson correlation analysis, Student *t* test, and analysis of variance. For nominal demographics (sex and nationality), variables were assumed to have metric characteristics and were recoded as follows: male = 1, female = 2; European = 1, non-European = 2. A negative correlation meant that there was a tendency for the demographic with a low value on one variable to have a high value on the other variable. A *P* value of less than .05 was considered statistically significant.

Results

A total of 216 trainees in the Irish Basic Surgical Training program participated in the study. All training centers in the country were represented. The group included 114 first-year trainees and 102 second-year trainees. There were 146 men and 70 women, 127 European nationals (including 105 Irish) and 89 non-Europeans. The mean age was 29.9 years, with non-European trainees being older on average (32.2 vs 28.3 y). Demographic details are summarized in Table 1.

Table 1 Demographic details of surgical trainees

Demographics	n (%)
Sex	
Male	146 (68)
Female	70 (32)
Nationality	
European	127 (59)
Male	71 (56)
Female	56 (44)
Non-European	89 (41)
Male	75 (84)
Female	14 (16)
Year of surgical training	
Year 1	114 (53)
Male	79 (69)
Female	35 (31)
European	61 (54)
Non-European	53 (46)
Year 2	102 (47)
Male	67 (66)
Female	35 (34)
European	66 (65)
Non-European	36 (35)

All scores were marked out of 100. Before the assessment, SPS were returned by 211 residents and the mean score was 70.9 (standard deviation, 14.1). These scores were compared with OAS. The correlation between OAS and SPS only weakly was positive (as shown by the gently upward sloping line and the broad scatter of marks around it) and was not statistically significant (Fig. 1; $r = .127$; $P =$ not significant).

After the test, SAS were returned by 195 trainees, with a mean of 69.1 (standard deviation, 14.5). These scores showed a moderate and statistically significant degree of correlation with the OAS (Fig. 2A; $r = .34$; $P < .001$), meaning that those residents who rated themselves highest after the assessment also were rated the highest by the faculty. A Bland and Altman[15] plot was used to assess the agreement between SAS and OAS. The mean difference between self-scoring and objective scoring was 8.9, with most candidates marking themselves lower than marked by examiners (Fig. 2B).

To evaluate the impact of demographic variables on self-prediction and self-assessment accuracy, the difference between these scores and the objectively assessed score was calculated for each individual (SPS-OAS and SAS-OAS). The distribution of SPS-OAS and SAS-OAS is shown in Fig. 3A and B, respectively. Looking at the group as a whole, most surgical trainees underestimated their competency. Mean SPS-OAS and SAS-OAS were compared for sex, year of training, and nationality (Table 2). Table 3

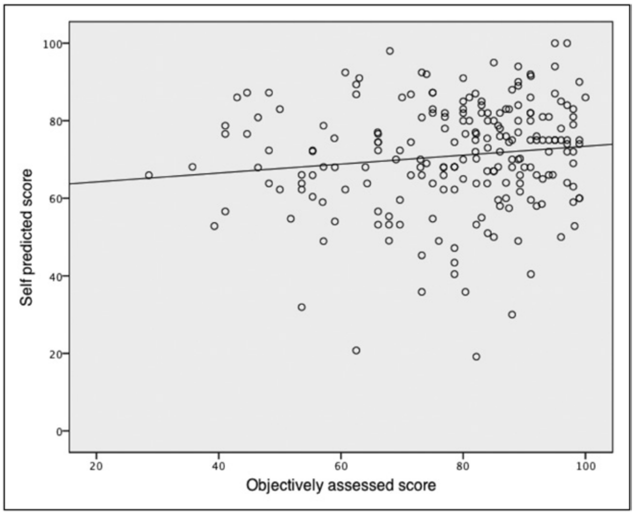

Figure 1 Correlation between objective (horizontal axis) and self-predicted (vertical axis) scores was only weakly positive, as represented by the gently sloping line and broadly scattered scores ($r = .127$; P = not specified).

summarizes relationships between resident demographic variables and each of the 2 accuracy scores (SPS-OAS and SAS-OAS). Because multiple dependents were being examined, analysis of variance also was performed and these results are summarized in Table 4.

There was no significant difference between the mean objective scores for male and female residents (77.58 and 79.56, respectively; $P = .24$). As shown by the negative correlation coefficients (Table 3), female trainees (coded with the lower value) underestimated scores more than their male counterparts. As illustrated in Fig. 4, this observation held true both before and after the assessment. This gender disparity was not significant, however, when other demographic variables were adjusted for during multivariate analysis (Table 4).

Significant differences in self-assessment accuracy depending on nationality, age, and year of training were identified on multivariate analysis (Table 4). European trainees tended to underestimate their scores to a higher degree than their non-European counterparts. Both age and year of training were associated with more accurate self-prediction and self-assessment ability. Second-year residents showed better self-prediction ability than first-year residents.

Comments

Self-awareness is recognized as an important personal trait in surgery and the current study examined 2 aspects of this: self-prediction and self-assessment accuracy. A significant relationship was found between residents' post-test self-assessed and objectively assessed scores but not between residents' prediction scores and actual objective scores. Although most trainees scored themselves lower than objective faculty examiners, there was low to moderate correlation between the 2 sets of marks ($r = .34$) (ie, trainees who gave themselves a higher score tended to receive a higher score and vice versa).

Almost 150 years ago, Charles Darwin (1871) observed that "ignorance more frequently begets confidence than does knowledge." This controversial concept—that when people make erroneous decisions, their incompetence deprives them of the ability to realize it—continues to be studied. In their 1999 article, "Unskilled and unaware of it," Kruger and Dunning[18] showed that college students who performed in the bottom quartile in examinations provided the most consistently inaccurate and inflated self-assessments. Conversely, the high performers were inclined to slightly underestimate their performance. Several studies in the psychology literature have replicated these findings[19–21] and, more recently, medical educators also have begun to consider this observation.[22,23] In a review of 20 comparisons between self-assessment and external assessment, Davis et al[24] found the worst accuracy in self-assessment among physicians who were the least skilled and those who were the most confident.

Previous studies in self-assessment of technical skill have shown at best a moderate association between self-assessed and observer-assessed scores of technical skill in surgery.[10,25–27] Ward et al[25] reported a moderate correlation ($r = .50$) between expert observer and resident self-evaluation after performance of a laparoscopic Nissen fundoplication on a pig. The level of association observed in this group was higher than that seen in our cohort and may be accounted for by the fact that the residents involved in the study were of a more senior level than the first- and second-year trainees that we examined. It has been suggested that improved self-assessment ability in more senior trainees may be the result of cumulative exposure to a greater number of cases with which they can compare their own skills.[25] Low levels of agreement between examiner and resident have been shown by Moorthy et al[27] among the most junior trainees included in their self-assessment study of a simulated saphenofemoral ligation. A number of strategies have been proposed to improve self-assessment accuracy. These include reviewing videotapes of one's own performance or of benchmark performances, as well as repetition of the assessed task.[10,25]

Trainees in our study consistently marked themselves lower than faculty observers. A similar observation was made by Mandel et al,[26] who asked a group of obstetrics and gynecology residents to score themselves after completion of an OSATS assessment of laparoscopic and open procedures. In contrast to this, a study of exit-level vascular surgery residents found that self-assessments exceeded grades awarded by examiners.[16] The investigators reasoned that residents at the end of their training may be reluctant to mark themselves in a manner that would suggest that they were not competent to perform the procedure in question. Similar findings were observed by Sidhu et al,[28] who examined self-assessment accuracy in a group of practicing

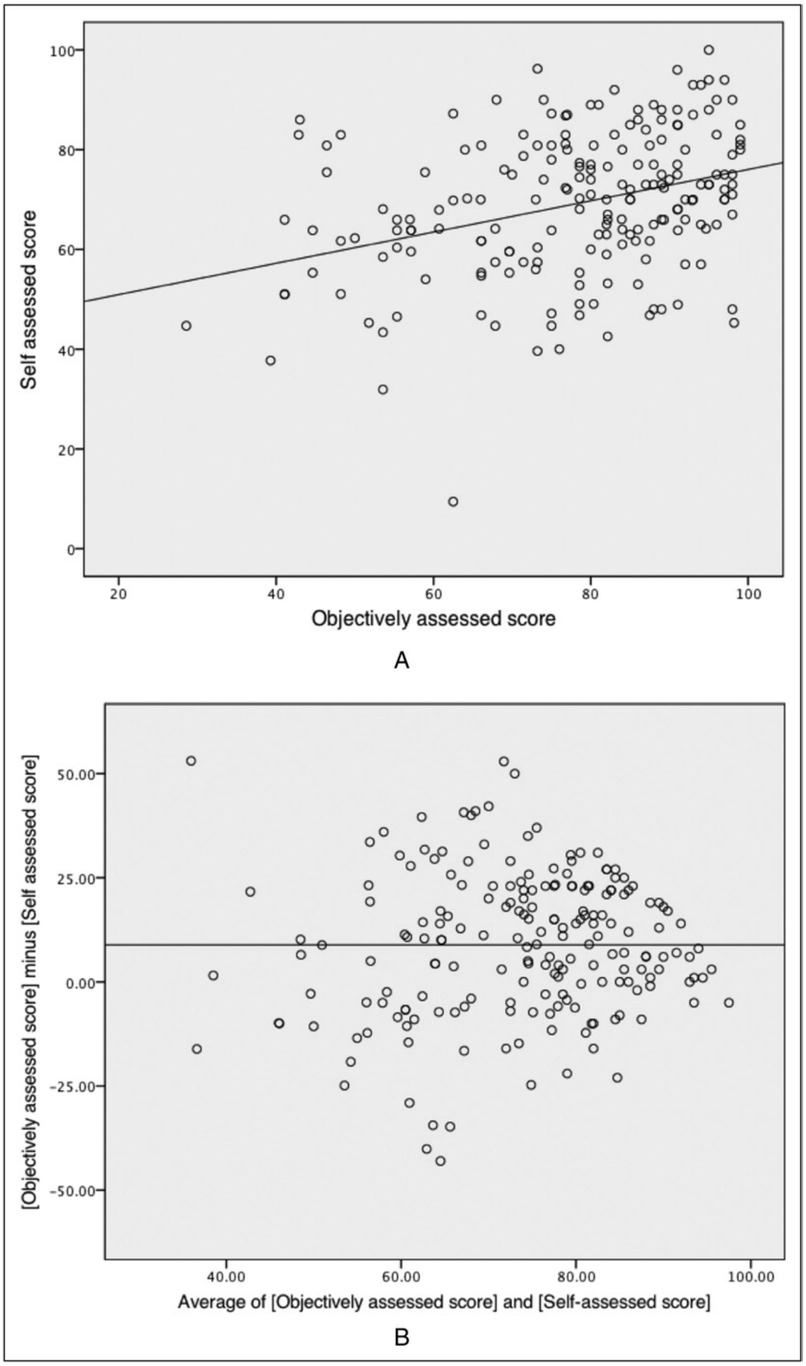

Figure 2 (A) Correlation between objective (horizontal axis) and self-assessed (vertical axis) scores was moderately positive and statistically significant ($r = .34$; $P < .001$). That is, those residents who rated themselves highest were rated highest by faculty. (B) Bland-Altman[15] plot further exploring this association, with clustering of scores above zero showing that faculty marked residents an average of 8.9 points (horizontal line) above the score they awarded themselves.

surgeons attending a laparoscopic cholecystectomy course. Both underestimation and overestimation of one's own ability are concerning traits in surgery. With the current emphasis on self-directed learning, junior surgeons who incorrectly estimate their ability may compromise the effectiveness of their training and this may impact on patient safety. Furthermore, accurate self-assessment is crucial for continuing professional development, which is now considered a core component of surgical practice.

Few medical education studies to date have considered demographic variables in examining self-assessment ability. Here, we report that older age, more senior year of training, and non-European nationality were predictive of more accurate self-prediction and self-assessment ability in junior surgical trainees. The association with year of training supports the conclusions of previous studies, which reported improved self-assessment accuracy with seniority.[25,27] No improvement with seniority was observed in a group of psychiatry residents, perhaps indicating that there is more informal feedback of technical skills/general competence in surgery than in other medical disciplines.[29]

Previous research in medical professionals has not shown an association between age and self-assessment accuracy.[29,30] Our finding may indicate that in the context of surgical training, younger trainees may require more positive feedback when they perform tasks well to become more accurate in their self-assessments.

The Irish Basic Surgical Training program delivers training to residents from a wide variety of backgrounds. Most trainees were European (127/216 [59%]) and, of these, 105 were Irish. The remainder came from outside of Europe, the majority were from Nigeria, Sudan, Malaysia, Mauritius, and Pakistan. The male majority was more significant in the non-European group, which is consistent with World Health Organization global health observatory figures, which show that men continue to dominate the medical profession in most of these countries.[31] It is not surprising that cultural differences would account for variation in self-assessment accuracy. In the present study, European trainees were found to underestimate their competence significantly in comparison with their non-European counterparts. Before undertaking this work, it might have been expected that a training program in Ireland naturally would be more suited to Irish medical graduates who are familiar with the cultural nuances of the health care system and the way in which training is delivered. The finding that local trainees are in fact less well able to self-assess needs to be addressed. Again, careful consideration should be given to the way in which surgical education, assessment, and feedback is delivered. A one-size-fits-all approach may not be appropriate.

Although women tended to underestimate their ability both before and after assessment, this was not found to be statistically significant at multivariate analysis. That is, although women did underestimate their competency, the

Table 2 Self-assessment accuracy according to demographic variables: sex, year of training, and nationality

Variable	SPS-OAS	SAS-OAS
Male sex	−5.08 ± 1.62*	−6.43 ± 1.55
Female sex	−12.37 ± 2.36	−13.5 ± 1.95
Year 1	−10.98 ± 1.49	−11.79 ± 1.45
Year 2	−3.33 ± 2.3	−5.67 ± 2.01
European	−12.92 ± 1.62	−14.32 ± 1.44
Non-European	.31 ± 2.06	−.61 ± 1.86

Self-assessment accuracy was calculated as differences in mean SPS/SAS and OAS for each demographic variable.

*Data are expressed as mean ± standard error of the mean.

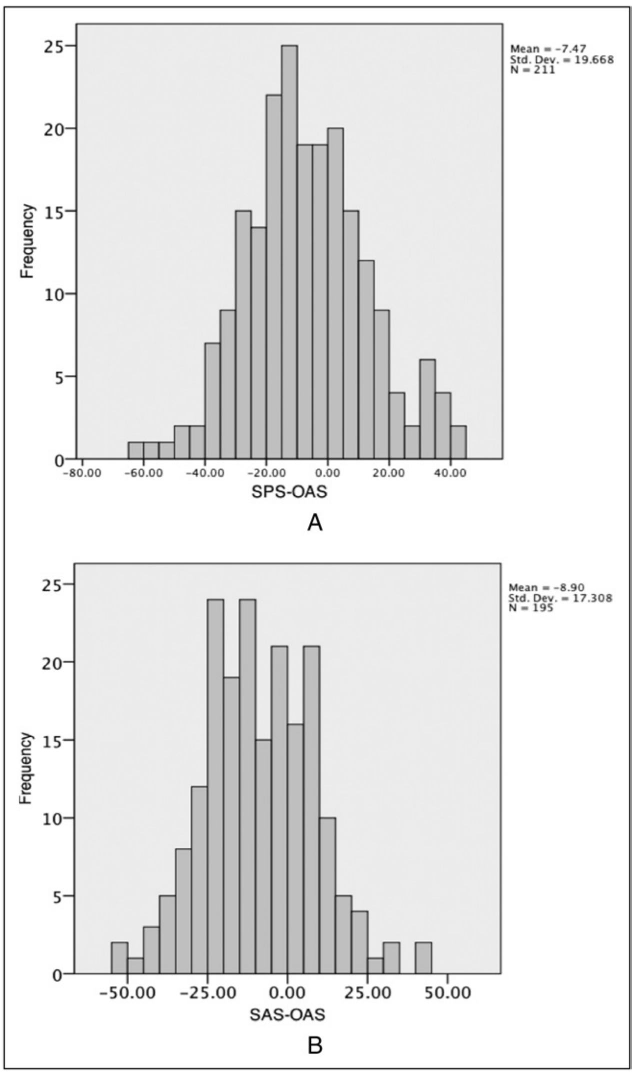

Figure 3 (A) Accuracy of self-prediction, calculated as the difference between residents' SPS and their objectively assessed score as awarded by faculty (OAS). Positive values mean that the objective score was higher, negative values mean that the self-predicted score was higher. Most residents underestimated their competency before undertaking the assessment. (B) Accuracy of self-assessment, calculated as the difference between residents' SAS and their OAS. Positive values mean that the objective score was higher, negative values mean that the self-assessed score was higher. Again, most residents underestimated their competency after having completed the assessment.

Table 3 Relationships between resident demographic variables and accuracy scores

Variable	Pearson correlation coefficient	*P* value
SPS-OAS vs sex	−.174	.01
SAS-OAS vs sex	−.195	.006
SPS-OAS vs age	.288	<.001
SAS-OAS vs age	.319	<.001
SPS-OAS vs nationality	.332	<.001
SAS-OAS vs nationality	.388	<.001
SPS-OAS vs year of training	.194	.005
SAS-OAS vs year of training	.177	.013

Nominal variables were coded as follows: male = 1, female = 2; European = 1, non-European = 2.

Table 4 Analysis of variance of self-prediction accuracy (SPS-OAS) and self-assessment accuracy (SAS-OAS) in relation to demographic variables

	SPS-SAS		SAS-OAS	
Variable	Coefficient of variance	*P* value	Coefficient of variance	*P* value
Sex	−2.030	.475	−1.589	.531
Age	.670	.033*	.625	.025*
Nationality	11.130	<.001*	11.534	<.001*
Year of training	9.001	<.001*	7.050	.002*

Nominal variables were coded as follows: male = 1, female = 2; European = 1, Non-European = 2.
*Statistically significant at a level of $P < .05$.

trend cannot be ascribed to gender. In a similar study, Minter et al[32] examined gender differences in surgical residents' self-assessment across a series of clinical competencies (including technical skill). A trend toward a greater degree of underestimation by female residents was observed but was not statistically significant. In a study of medical students, Lind et al[33] showed that women on surgical rotations significantly underestimated their abilities. The established standard of surgery as a male-dominated profession may have an influence on junior female trainees' low confidence in their technical skill. If inaccurate self-perceptions do exist in surgical trainees based on gender, they need to be addressed. Studies have shown that the inability to see oneself as competent in a given occupation may affect career choice, performance, and persistence within that domain, and may in part account for under-representation of women in male-dominated fields.[34,35] In our cohort, female trainees formed a significant minority (32% of basic surgical trainees) and this was consistent with surgical training programs internationally.[36] Although it is unfortunate that women underestimate their performance, it is useful to know that this tendency is not based on their gender.

The results of this study must be interpreted within certain limitations. First, it is possible that more senior residents may have had enough experience with the assessment process to show improvement over their first-year counterparts. Second, there were additional factors that were not analyzed in this study that may have influenced self-assessment ability. These include musicality, athleticism, or participation in other competitive processes. The influence of extracurricular activities on surgical technical skill is an ongoing area of research in our department,[37] and it would be interesting to explore the association that such pursuits may have with self-assessment ability. Finally, this study was conducted in a skills laboratory setting and this may be a limitation to the generalizability of the findings. Furthermore, it is possible that the results obtained were influenced by the relatively stressful nature of the day that the survey was performed. These limitations could be addressed by repeating the study in an operating room environment.

Conclusions

The Irish Basic Surgical Training program affords a unique opportunity to examine self-assessment accuracy in a large and diverse group of junior surgical trainees. Although other studies are limited by the small size of most surgical training programs, the large number in this cohort has allowed significant conclusions to be reached and demographic variables to be included in the analysis. Direct observation with a specific checklist of marking criteria is considered the most reliable method of assessing surgical technical skill and the OSATS approach has been shown to have a high degree of reliability and validity.[14,38] Residents' OSATS results therefore were considered a reliable gold

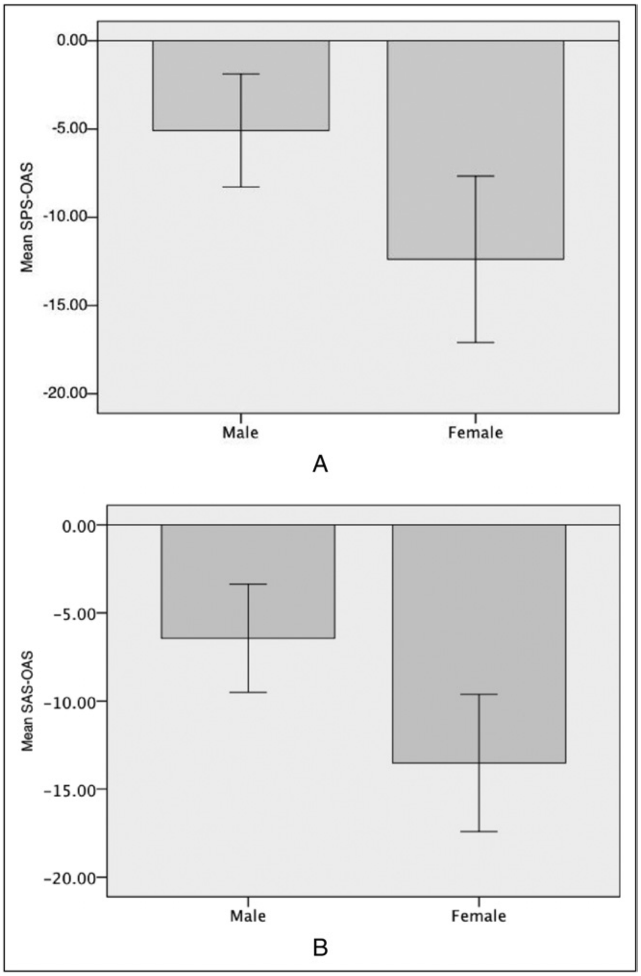

Figure 4 Difference in self prediction and self-assessment accuracy based on gender, displayed as the difference between residents' (A) SPS and their (B) objectively assessed score (SPS or SAS–OAS). Male and female residents underestimated their competency both before and after the assessment, although women did so to a greater degree.

standard with which their self-predicted and self-assessed scores could be compared. The process used to examine self-assessment ability involved minimal time and expense and was administered easily by faculty on the day of trainees' clinical skills assessment.

Our ability to deliver effective surgical education is contingent on our understanding of potential barriers to learning and the differences among trainees that affect their ability to develop and progress in their careers. We have identified several demographic variables that are associated with more accurate performance prediction and self-assessment of technical skill, specifically older age, non-European nationality, and more experience in training. More important, however, is the ability to anticipate the trainees who need assistance in developing better self-awareness of their technical skills and to examine methods for improving this important attribute.

References

1. Kohn LT, Corrigan JM, Donaldson MS. To Err Is Human. Washington, DC: National Academies Press; 2000.
2. Working Group on Specialist Medical Training. Hospital Doctors: Training for the Future. London: Department of Health; 1993.
3. Chung RS. How much time do surgical residents need to learn operative surgery? Am J Surg 2005;190:351–3.
4. Feanny MA, Scott BG, Mattox KL, et al. Impact of the 80-hour work week on resident emergency operative experience. Am J Surg 2005; 190:947–9.
5. Carlin AM, Gasevic E, Shepard AD. Effect of the 80-hour work week on resident operative experience in general surgery. Am J Surg 2007; 193:326–9; discussion, 329–30.
6. Chikwe J, de Souza AC, Pepper JR. No time to train the surgeons. BMJ 2004;328:418–9.
7. Blum AB, Raiszadeh F, Shea S, et al. US public opinion regarding proposed limits on resident physician work hours. BMC Med 2010;8:33.
8. Johna S. Limitations in resident work hours: are we preaching to the choir? That is the question! Arch Surg 2011;146:11.
9. Grantcharov TP, Reznick RK. Teaching procedural skills. BMJ 2008; 336:1129–31.
10. MacDonald J, Williams RG, Rogers DA. Self-assessment in simulation-based surgical skills training. Am J Surg 2003;185:319–22.
11. Accreditation Council for Graduate Medical Education. ACGME program requirements for graduate medical education in surgery. Available from: http://www.acgme.org/acWebsite/downloads/RRC_progReq/440_general_surgery_01012008_u08102008.pdf. Accessed: February 3, 2011.
12. Cahan MA, Larkin AC, Starr S, et al. A human factors curriculum for surgical clerkship students. Arch Surg 2010;145:1151–7.
13. Traynor O. SE11 development of a personal Skills Programme for Surgical Trainees. Aust N Z Surg 2007;77:A79.
14. Reznick RK. Teaching and testing technical skills. Am J Surg 1993; 165:358–61.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
16. Pandey VA, Wolfe JH, Black SA, et al. Self-assessment of technical skill in surgery: the need for expert feedback. Ann R Coll Surg Engl 2008;90:286–90.
17. Evans AW, Leeson RM, Petrie A. Reliability of peer and self-assessment scores compared with trainers' scores following third molar surgery. Med Educ 2007;41:866–72.
18. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol 1999;77:1121–34.
19. Krueger J, Mueller RA. Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. J Pers Soc Psychol 2002;82:180–8.
20. Dunning D, Griffin DW, Milojkovic JD, et al. The overconfidence effect in social prediction. J Pers Soc Psychol 1990;58:568–81.
21. Vallone RP, Griffin DW, Lin S, et al. Overconfident prediction of future actions and outcomes by self and others. J Pers Soc Psychol 1990;58:582–92.
22. Hodges B, Regehr G, Martin D. Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. Acad Med 2001;76(Suppl):S87–9.
23. Eva KW, Cunnington JP, 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 Theory Pract 2004;9:211–24.
24. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA 2006;296:1094–102.
25. Ward M, MacRae H, Schlachta C, et al. Resident self-assessment of operative performance. Am J Surg 2003;185:521–4.
26. Mandel LS, Goff BA, Lentz GM. Self-assessment of resident surgical skills: is it feasible? Am J Obstet Gynecol 2005;193:1817–22.
27. Moorthy K, Munz Y, Adams S, et al. Self-assessment of performance among surgical trainees during simulated procedures in a simulated operating theater. Am J Surg 2006;192:114–8.
28. Sidhu RS, Vikis E, Cheifetz R, et al. Self-assessment during a 2-day laparoscopic colectomy course: can surgeons judge how well they are learning new skills? Am J Surg 2006;191:677–81.
29. Lynn DJ, Holzer C, O'Neill P. Relationships between self-assessment skills, test performance, and demographic variables in psychiatry residents. Adv Health Sci Educ Theory Pract 2006;11:51–60.
30. Leopold SS, Morgan HD, Kadel NJ, et al. Impact of educational intervention on confidence and competence in the performance of a simple surgical task. J Bone Joint Surg Am 2005;87:1031–7.
31. World Health Organization. Male to female ratio of physicians density per 1000 population, latest available year. Available from: http://www.who.int/gho/health_workforce/physicians_density_gender/en/index.html. Accessed: February 12, 2012.
32. Minter RM, Gruppen LD, Napolitano KS, et al. Gender differences in the self-assessment of surgical residents. Am J Surg 2005;189:647–50.
33. Lind DS, Rekkas S, Bui V, et al. Competency-based student self-assessment on a surgery rotation. J Surg Res 2002;105:31–4.
34. Beyer S. Gender differences in the accuracy of self-evaluations of performance. J Pers Soc Psychol 1990;59:960–70.
35. Beyer S, Bowden E. Gender differences in self-perceptions: convergent evidence from three measures of accuracy and bias. Pers Soc Psychol Bull 1997;23:157–72.
36. Leadley J. Women in US Medicine Statistics and Benchmarking Report. Washington, DC: Association of American Medical Colleges; 2009.
37. Boyle E, Kennedy AM, Traynor O, et al. Training surgical skills using nonsurgical tasks—can Nintendo Wii™ improve surgical performance? J Surg Educ 2011;68:148–54.
38. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 1997;84:273–8.