Obstacles Affecting the Implementation of the O-SCORE for Assessment of Orthopedic Surgical Skills Competency

Gabriella E. Ode, MD, Joseph Stewart Buck, MD, Meghan Wally, MSPH, Brian P Scannell, MD, and Joshua C Patt, MD

Department of Orthopaedic Surgery, Carolinas Medical Center, Charlotte, North Carolina

OBJECTIVES: There is a need for meaningful and reliable measures of surgical competency in residency education. The goal of the current study is to incorporate the Ottawa Surgical Competency Operating Room Evaluation (O−SCORE) into the process of resident evaluation at our institution and to assess the feasibility and effectiveness of its use through a web−based platform.

DESIGN: This is a feasibility study that prospectively assesses the implementation of a web-based O-SCORE at our institution. Over a 16−week period, 19 orthopedic surgery residents (PGY2−PGY5) participated in a quality improvement study, which involved collecting 2 feedback forms per week. Each form consisted of a resident form and a linked attending form. At the conclusion of the 16−week trial period, residents and faculty members were asked to complete a survey about their perceptions of the O−SCORE program.

SETTING: An academic medical center.

PARTICIPANTS: The study included only residents in postgraduate training years (PGY) 2 through 5 (n = 20) and attendings (n = 37).

RESULTS: During the 16−week study period, 608 resident surveys were requested for the 19 participating residents, of which 404 surveys (66.5%) were completed. Faculty completed 207 of 326 surveys for an overall compliance rate of 63.5%. The O−SCORE was able to significantly differentiate between all training years ($p < 0.0001$) with the exception of PGY3 residents when compared to PGY4 residents. Overall, residents and faculty found the program valuable and feasible. Resident and faculty perception of the value of the O−SCORE correlated with compliance rate of the O−SCORE surveys.

CONCLUSIONS: This study demonstrates that implementation of an immediate feedback program utilizing an electronic platform is achievable and offers reproducible construct validity. However, issues affecting compliance among both residents and faculty physicians must temper optimism for the program and should be systematically addressed to allow for successful implementation. (J Surg Ed 76:881−892. © 2018 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: Ottawa Surgical Competency Operating Room Evaluation, Orthopedic surgery, Competency, Education

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

INTRODUCTION

The resident competency evaluation process in orthopedic surgery is evolving. Stimulus for change comes from the changing landscape of residency education. Current ACGME regulations for evaluation of both operative and clinical milestones that were set forth by the implementation of the Next Accreditation System in 2012 have increased the number of evaluations done, but this has not led to improvement in direct resident feedback on individual competency. Recent studies have reported a significant disparity between residents and faculty satisfaction with postoperative feedback with residents reporting relatively higher rates of dissatisfaction with the quality and timeliness of their surgical feedback.[1,2] An internal survey among the residents of our program reported that 63% of residents were not satisfied with their feedback

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Correspondence: Inquiries to Gabriella E. Ode, MD, Department of Orthopaedic surgery, Carolinas Medical Center, 1000 Blythe Blvd Charlotte, NC; e-mail: gabriellaode@gmail.com

  1931-7204/$30.00 https://doi.org/10.1016/j.jsurg.2018.11.002

on operative performance specifically. Our residents most frequently cited a lack of case-specific feedback and the infrequency of feedback as their sources of dissatisfaction. In the current graduate medical education literature, there has been a growing trend toward the utilization of entrustability scales for assessment of procedural competency among surgical residents[3-5] Entrustability scales are defined as behaviorally anchored ordinal scales based on progression to competence, which are based on the clinical judgment of degree of supervision required.[3] For example, the more supervision desired by the attending indicates a lower level of entrustability.

The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE)[6] is of a recently developed entrustability scale that assesses a trainee's ability to independently perform a surgical procedure ranging on a scale from limited entrustability ("requires complete hands on guidance") to complete entrustability ("complete independence performing task"). Adopters of entrustability scales for surgical education like the O-SCORE argue that these scales are superior to the current operative case log system in which presence is equated with competence. In contrast, these plain-language scale focus on clinically relevent performance and can better stratify trainees and identify areas for improvement.[6] Furthermore, these scales may also fulfill the needs of residents who desire more frequent and specific feedback by deliberately aligning with day-to-day assessments of competency and independence in the setting of clinical education. Previous work demonstrated that the O-SCORE is a valid and reliable tool for the assessment of surgical competency among orthopedic and general surgery residents[6]; however, the feasibility of the O-SCORE as an assessment tool has yet to be externally evaluated. The purpose of the current study is to assess the incorporation of an entrustability scale evaluation form for operative performance (O-SCORE) at our institution and to assess the feasibility and effectiveness of its use on a web-based platform.

MATERIALS AND METHODS

A web-based version of the O-SCORE was developed at our institution using RedCap (Research Electronic Data Capture), a secure web-based electronic data capture tool for research studies.[7] Over a 16-week period between August and November 2016, 19 orthopedic surgery residents (PGY2-PGY5) participated in a quality improvement study, which involved the completion of 2 O-SCORE forms per week. The study included only residents in postgraduate training years (PGY) 2 to 5 (n = 20). First year residents (PGY1) were excluded from participation. Participation was voluntary. One resident declined to participate in the pilot study, leaving 19 residents who participated for the entire 16-week period. Prior to initiation of the study, residents were given a brief presentation explaining the O-SCORE and their individual responsibility to initiate the feedback encounter. Participants were given a list of 19 common orthopaedic procedures that were recommended for evaluation but any orthopedic surgical procedure could be evaluated (See Appendix 1).

Each postoperative feedback evaluation consisted of 2 parts: a resident component that specified the procedure performed, the name of the supervising attending, and the level of resident involvement in the case (Fig. 1A). The second portion was an attending form that consisted of the O-SCORE immediate postoperative feedback form (Fig. 1B-D). At the start of each week, residents received 2 blank resident forms, and 2 blank corresponding attending forms via electronic mail (e-mail). Resident involvement was defined via ACGME case log guidelines[8] as either Level 1 (primary or supervising resident surgeon role) or Level 2 (assisting resident surgeon role). Subjects could be excused from completing a given form if they were unable to complete the form that week due to call/night float responsibilities or unavailability of the resident or assigned attendings physician for surgical procedures that week (e.g., away for vacation or conference). Completion of the resident form would auto-populate portions of the corresponding attending form (resident name, date, time, and surgical procedure, which the resident then sent by e-mail to the supervising attending for completion). An automatically generated e-mail reminder to complete the resident form was sent 4 days after the initial e-mail if the form had not been completed. No further reminders were sent. Automatic e-mail notification of completion of the attending form was sent to residents in real time. The RedCap platform was not capable of sending reminders directly to attending physicians, so residents were responsible for verbally reminding their attending to complete their assigned O-SCORE form if they were delinquent. Compliance with form completion by residents and faculty were recorded.

Since 1 study goal was to assess the baseline feasibility and utility of the O-SCORE form within our program, no further interventions (punitive or compensatory) were used to encourage completion of the resident and faculty forms each week. Residents were provided a compilation of their scores at the 8-week point in the study and at the end of the 16-week data collection period. Residents could not see the scores of their peers nor could faculty members obtain access to the scores of any resident after completion of their assigned O-SCORE. At the end of the 16-week feedback period, residents and faculty surgeons participated in separate focus group discussions about the O-SCORE form and the feedback program and completed a survey about their overall perception of the feedback program.

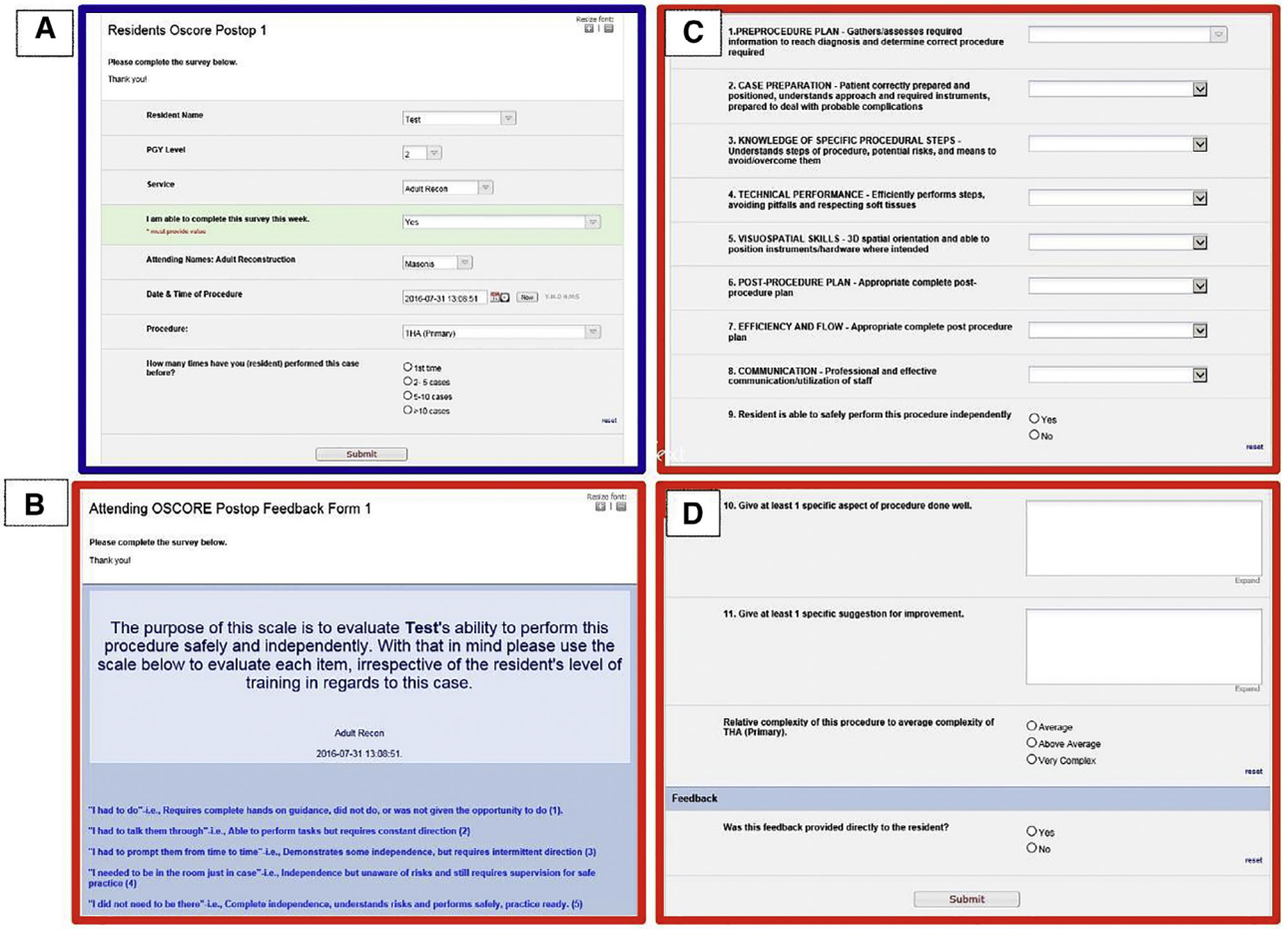

FIGURE 1. OSCORE resident and attending web-based forms.

Analysis

Data were compiled in the REDCap database and exported to Statistical Analysis System for analysis. Descriptive statistics were used to report the primary outcome measure of compliance rate. Surveys that were completed by the resident and sent to the attending and completed surveys indicating excused absences were considered compliant for the purposes of this analysis. An intent-to-treat approach was used; therefore, all surveys requested of residents were included in the denominator. Construct validity of the O-SCORE and statistical differences in average score by PGY level were assessed through ANOVA. Finally, a descriptive analysis was performed to report on requested procedures and resident and faculty feedback.

RESULTS

Resident Compliance

During the 16-week study period, 608 resident surveys were expected for the 19 participating residents, of which 404 surveys (66.5%) were completed. After excluding surveys that indicated excused absences for call/night float responsibilities or unavailable resident/faculty, there were 326 survey forms that were sent to faculty for feedback (53.6%). This equated to 1.1 surveys per week that were distributed by residents to 37/39 faculty members. Two faculty members who were not assigned to a participating resident on their rotation during the OSCORE collection period did not receive a request for an O-SCORE form and were excluded from analysis. Resident compliance by program year is summarized in Figure 2. Faculty completed 207 of 326 surveys for an overall compliance rate of 63.5%. In 77.8% of cases, immediate verbal feedback at the time of the O-SCORE survey completion was provided to residents (n = 161/207).

Mean Scores

The mean scores for each O-SCORE category by program year are summarized in Figure 3. The O-SCORE demonstrated good construct validity: it was able to

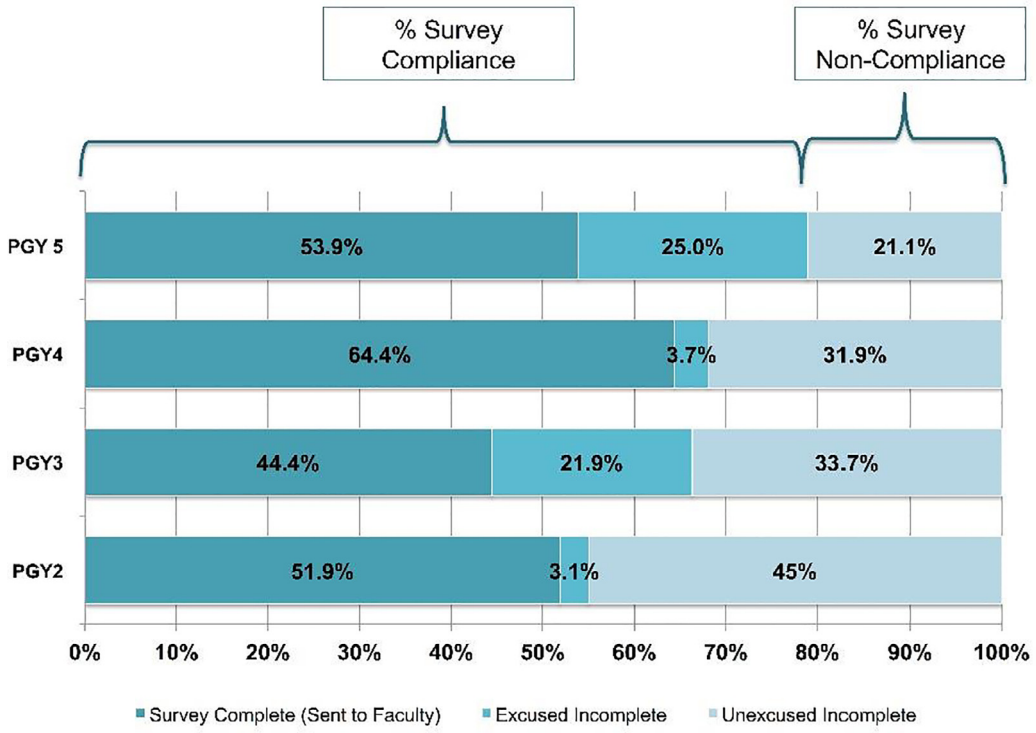

FIGURE 2. Compliance rate by program year (PGY2-PGY-5).

significantly differentiate between all training years ($p < 0.0001$) with the exception of PGY3 residents when compared to PGY4 residents. In terms of the descriptive question regarding a residents' ability to safely perform a procedure, over 80% of PGY5 residents were considered as safe to perform a procedure independently compared to only 24% of PGY2 residents (Fig. 4).

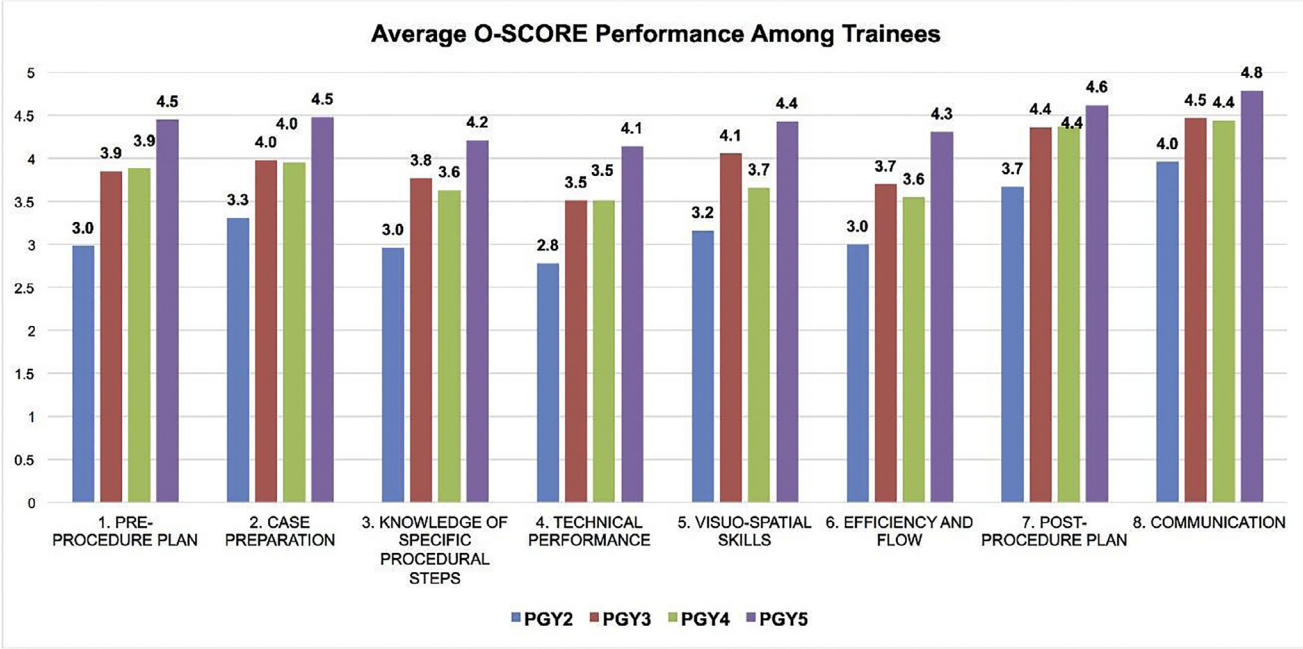

FIGURE 3. Average O-SCORE performance by training year.

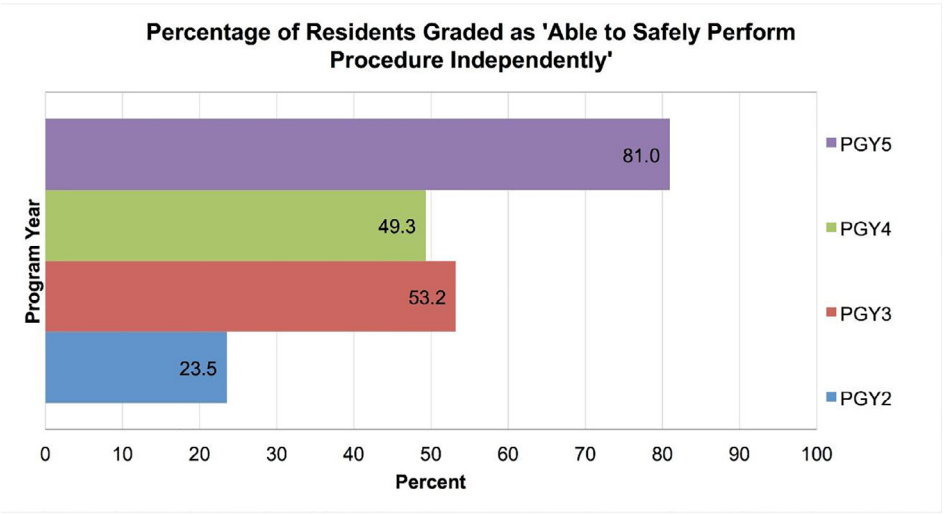

FIGURE 4. Residents graded by ability to safely perform procedure independently.

Procedures

There were 326 requested O-SCORE surveys, of which 312 appropriately reported the surgical procedure (96%). The most frequently requested procedures were total knee arthroplasty (9%), total hip arthroplasty (7.4%), operative fixation of diaphyseal tibia/femur fractures (7.1%), operative fixation of ankle fractures (5.5%), operative fixation of distal radius fractures (4.2%), and percutaneous fixation of pediatric supracondylar humerus fractures (4.2%), all of which are ACGME-recognized procedures through either the Milestone Project or Case Log Minimums. Overall, 58% of the requested evaluations were of an ACGME Milestone recognized procedure while non-ACGME Milestone recognized procedures made up 6% of cases. In the remaining 36% of cases, an "other" procedure was reported which indicated a surgical procedure outside of the list of 19 recommended procedures. The 2 most common "other" procedures were excision of soft tissue mass (n = 11; 3.5%) and operative fixation of forearm fracture (n = 8; 2.5%).

Resident/Faculty Feedback

At the conclusion of the 16-week trial period, residents and faculty members were asked to complete a survey about their perceptions of the O-SCORE program. All 19/19 residents and 27/39 faculty members completed the final survey.

A slight majority of residents (58%) and faculty (56%) found the O-SCORE feedback program to be valuable for resident training. Similarly, a majority of residents (58%) and faculty (78%) felt that 2 forms per week was an appropriate amount to elicit feedback. Regarding the O-SCORE questions, 70% of faculty described the O-SCORE questions as appropriate in length and in detail, compared to 58% of residents. Nearly 1/3 of residents found the form to be too long or overly detailed compared to 11% of faculty. Perceived value of the feedback program is illustrated in Figure 5. Resident and faculty perception of the value of the O-SCORE correlated with compliance rate of the O-SCORE surveys. When comparing the compliance rates of subjects based on their opinion regarding the value of the O-SCORE, increased compliance rates were reported among those who found the O-SCORE to be valuable compared to those who had a neutral opinion of the O-SCORE and those who did not find the O-SCORE to be valuable ($p = 0.0244$). Among residents, having a neutral opinion correlated with a significantly lower completion rate than having either a positive or negative opinion. There were statistically significant differences in compliance by perceived value ($p < 0.0001$). Those with a neutral opinion had lower completion rates than those with either negative or positive opinions. Themes reported during resident and faculty opinion polls regarding the feedback program are described in Tables 1 and 2, respectively. Overall, both faculty and residents most often endorsed the fact that the OSCORE survey encouraged real-time structured feedback as a benefit of this program. However, the most common complaint about the program by faculty was that questions were too vague, difficult to assess, or irrelevant, while residents most often complained about the technology being inconvenient or cumbersome.

DISCUSSION

There is a dire need for meaningful and reliable measures of surgical competency in residency education.

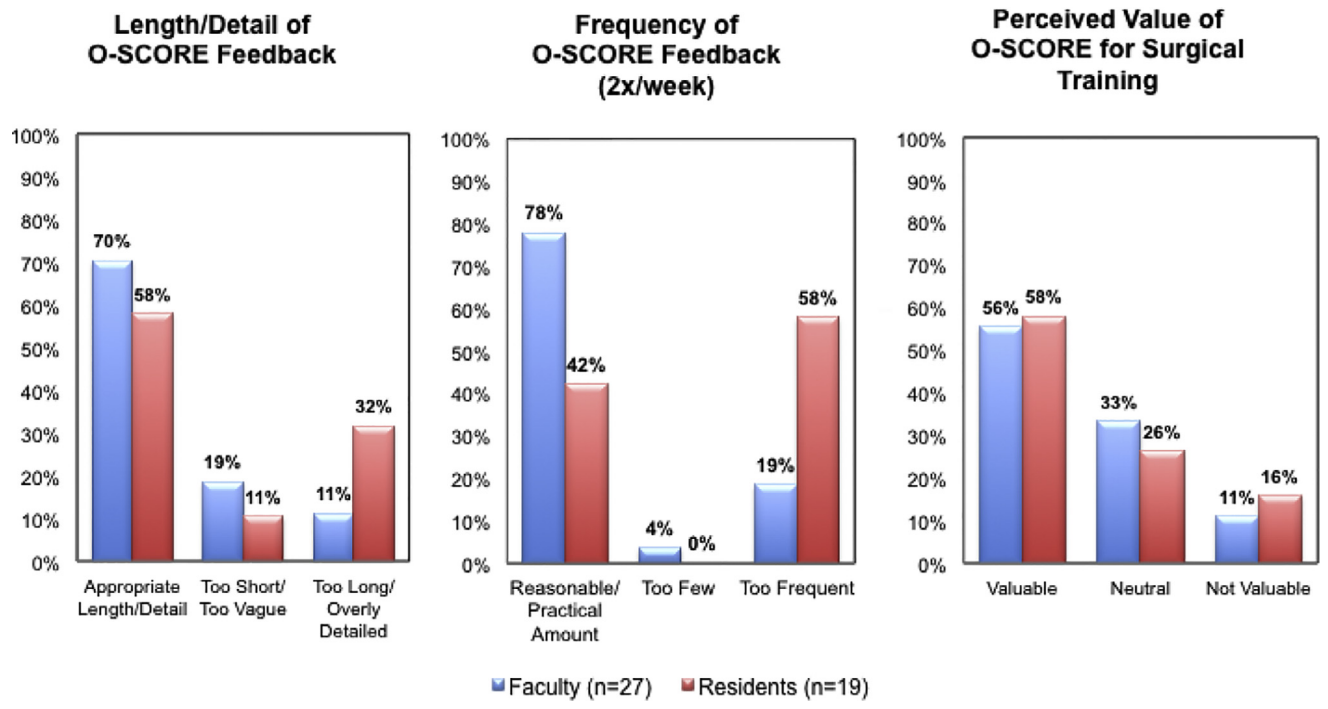

FIGURE 5. Resident and faculty perceptions of the O-SCORE feedback program.

There is a growing trend toward the implementation of immediate surgical skills assessment in surgical training. While the O-SCORE represents a recently validated form[6] utilized in orthopedic residency, several other studies have evaluated similar assessments. Cooney et al.,[4] designed a web-based assessment tool for plastic surgery residents called the Operative Entrustability Assessment, which was also based on an entrustability scale of operative autonomy. Residents were required to initiate a self-assessment of their performance on every operative case, which was then sent to their attending physician for further comment. Attendings were responsible for assigning the relevant milestone and grading the resident based on his/her performance. The authors reported that the Operative Entrustability Assessment demonstrated construct validity by effectively distinguishing novice from experienced residents.[4] Furthermore, the touted benefits of their system was that it increased immediate attending/trainee feedback and assessment transparency, informed end-of-rotation reviews and program-wide assessments, and enabled trainee self-monitoring so that trainees could address specific needs all in real-time. Cook et al.[9] designed and implemented a flexible feedback process in which residents initiated a postoperative feedback discussion and completed a Procedure Feedback Form (PFF) with their supervising attending surgeon. Residents were expected to initiate a feedback discussion and complete a PFF with their attending surgeons after selected cases. Residents were expected to complete 10 PFFs every 6 months and these were reviewed at the Core Competency Committee meetings and at the residents' subsequent meeting with their advisors. They found that the form was adaptable to a variety of cases and demonstrated good reliability between the resident and attending assessments. Wagner et al.[10] looked at an immediate feedback form and utilized a mobile platform to facilitate form completion. They noted that, while using their form, constructive feedback discussions between residents, and faculty members occurred regularly and more rapidly after each operation.

The O-SCORE uses entrustability language instead of a comparison scale of performance to peers as a means to avoid end aversion bias and allows for differentiation among different levels of training. Gofton et al.[6] found that residents were comfortable receiving lower O-SCOREs when the scale was worded in terms of entrustment. While the initial results of the O-SCORE appear promising, change within the culture of surgical training is difficult and compliance with the implementation of this surgical assessment tool has not been explicitly reported. The current study assessed the feasibility of the O-SCORE immediate feedback program for achieving reproducible surgical feedback among orthopedic residents in a single training program. Furthermore, it attempts to report some of the common obstacles that may affect compliance during the early period of implementation of such a program. Our study confirmed

TABLE 1. Resident Opinion Survey on the OSCORE Feedback Program

What Did You LIKE About the O-SCORE Survey and the Feedback Program?	
Aggregate Statements (Number of Participants)	**Example Comments**
Encouraged real-time structured feedback/dialogue (n = 10)	*"It provided a structured way to ask and receive detailed feedback from the attendings."* *"Initiated conversation/face to face feedback discussion between attendings and residents."*
Enriched learning/teaching experience (n = 3)	*"It forced the attending to have to objectively think about my learning experience. It also provided a bench mark for me to aim for."* *"Makes you think about performance on a daily basis."*
Nothing (n = 3)	
Other (n = 1)	*"Frequency. Twice a week was a good number."*
What Did You DISLIKE About the O-SCORE Survey and the FeedbackProgram?	
Inconvenient or cumbersome technology (n = 8)	• *"Email interface made it difficult. An app would be much more user friendly."* • *"Took extra time that is hard to find. I do not feel this form was helpful to facilitate feedback."* • *"Needing to email the form to attendings was a hassle and one more step likely to fail."*
Limited ability to review real-time feedback (n = 5)	• *"Feedback is not real time. Though it is encouraged that we actually discuss the feedback, it does not always happen. Real time feedback from the survey at least gives the resident an opportunity to further tend to their opportunities to improve."*
Questions too vague, difficult to assess or irrelevant (n = 5)	• *"If done correctly, the form takes significant amount of time to fill out. There are not always appropriate cases each week that would fit the requirements for these forms."* • *"The questions don't go well with my level."*
Felt uncomfortable asking for feedback (n = 4)	• *"It is easy for residents to cherry pick cases which they felt went well because 1) it's always nicer to hear positive feedback than negative feedback, and 2) no one wants their program administrators (eg program director) to see negative feedback when it's possible."* • *"Felt like I was intruding with some of the attendings."*
Too frequent (n = 2)	• *"Email interface made it difficult. An app would be much more user friendly."* • *"Took extra time that is hard to find."* • *"Needing to email the form to attendings was a hassle and one more step likely to fail."*

Gofton et al.'s results that the O-SCORE demonstrates good construct validity and can differentiate between training years, particularly between junior and senior residents.[6] However, the study also showed that compliance with use of the O-SCORE on a web-based platform was relatively modest among both residents (66.5%) and faculty (63.5%).

We uncovered several issues that may have affected study compliance. One issue was related to the frequency that residents wanted to receive feedback. The majority of residents (58%) thought that 2 feedback forms per week was too frequent while 78% of faculty thought this was an ideal amount. This discrepancy may be because residents were reminded electronically to complete their O-SCORE forms while the faculty were only verbally reminded by residents to complete the corresponding O-SCORE form. As part of the design of the e-mail-based platform, a resident who had not completed their O-SCORE forms for a given week would receive 4 e-mails per week (2 initial e-mails with that week's O-SCORE request and then 2 automatically generated reminder e-mails at the end of the week if the forms were not completed). The frequent e-mails may have led to "alarm fatigue," which may have led to disinterest in the program. In contrast, faculty members were not sent reminder e-mails and subsequently may not have experienced the same fatigue toward evaluation requests. Furthermore, since most residents worked with multiple

TABLE 2. Faculty Opinion Survey on the OSCORE Feedback Program

What Did You LIKE About the O-SCORE Survey and the Feedback Program?	
Aggregate Statements (Number of Participants)	**Example Comments**
Encouraged real-time structured feedback/dialogue (n = 11)	• *"The idea of feedback on a regular basis is important. This program stimulated me to discuss more with the residents on my service at a regular basis."* • *"Productive real-time feedback"* • *"The answer categories are descriptive and provide real case feedback about the resident's ability to function independently."*
Nothing (n = 3)	
Enriched learning/teaching experience (n = 2)	• *"I very much enjoyed when the resident sent it to me immediately after the case and we could sit down and I filled it out and discussed it with them. Much better feedback and gave me more insight into the resident."* • *"It strengthens case-based learning."*
Other (n = 6)	• *"Short"* • *"Objective"* • *"Asked appropriate questions"*
What Did You DISLIKE About the O-SCORE Survey and the Feedback Program?	
Questions too vague, difficult to assess or irrelevant) (n = 10)	• *"I feel we are trying to overly objectify what we all know. . .'Is a resident competent?' This could be simplified to a single question—did the resident complete the case at a level expected of their year of training. And if no, was it surgical skill, knowledge or preparation?"* • *"Fairly vague scores. I felt it was hard to judge where a resident should be rated. When a resident hasn't done a certain case with me, I walk them through how I do the case—how then do I rate that resident?"* • *"The choices aren't appropriate for about half of the questions. Answering that I didn't need to be in the room for the pre and post op plans is weird."* • *"Evaluation for evaluation's sake is not helpful. Coaching on skills done well and possibilities for future changes is a more appropriate method."* • *"Some questions are not perfect. For example, a second-year resident should not be able to do the majority of the spine cases on their own without supervision. I certainly understand the long-term goal of it however that is one of the shortcomings."* • *"The responses are too often non-sensical as they are applied in the same manner to resident at all levels. 1. Rarely do I feel that I can leave a resident alone to do a case 2. Rarely, do I think that I don't have something to add or impart that makes my presence in the case worthwhile. 3. Few residents beneath the PGY-5 level are ever able to do cases unsupervised - to the point where I could state that I didn't need to be there. We need to stratify the response by PGY level and better focus on the expectations at each level For this project—one size doesn't fit all."*
Nothing (n = 6)	
Poor compliance/delayed requests by residents (n = 4)	• *"Often I would receive the OSCORE days after the case. It must be sent and completed in a timely fashion. The responsibility needs to remain on the resident for this to work."* • *"I don't like that it is at the discretion of the resident to ask us to fill out."* • *"The choice of cases is very random and arbitrary—We need to formulate a list of cases per specialty to focus on. Residents were very inconsistent in sending me the link to complete these as well."*

(*continue on next page*)

TABLE 2 (*CONTINUED*)

What Did You DISLIKE About the O-SCORE Survey and the Feedback Program?	
Too frequent (n = 2)	• *"Decrease to one per week. This allows improvement before another score."*
Inconvenient or cumbersome technology (n = 1)	• *"Doing it on the resident's phone right after the case is not the best time. If we had the app on our phone and was alerted to fill it out, that would be better."*

faculty members during a given week, it is unlikely that faculty consistently received requests for greater than 2 forms per week and thus they may not have felt the burden of excessive forms each week. Previous literature has demonstrated a positive effect using electronic reminders to improve the rate of return for evaluations,[11,12] however the frequency of these reminders ranged from once at the end of a 9 or 12-month rotation[11] to once a day for a daily evaluation program. Therefore, in the context of this study, it is difficult to quantify the magnitude of the effect of the reminder e-mail intervals on faculty perception of the evaluation requests. One potential solution to the alarm fatigue would be to decrease the frequency of the reminder e-mails to a single e-mail per week or every 2 weeks while maintaining the frequency of the feedback forms (2 per week). Alternatively, given the resident dissatisfaction with the form frequency, another potential compromise for both residents and faculty would be decrease the form frequency to 1 form per week but create a platform that sent a reminder e-mail to faculty who had not completed their weekly form. This could mitigate the alarm fatigue experienced by the residents but also offer a structured reminder for faculty in order to increase compliance.

We noticed a trend toward lower survey compliance among junior residents compared to more senior residents. One explanation is that part of the natural maturation curve in residency training means senior residents are inherently more open to receiving written feedback than junior residents. Alternatively, one could argue that the significantly lower O-SCOREs among junior residents may have affected their desire to continue to seek feedback. Both positive and negative feedback may affect a learner's motivation to seek more feedback.[13] Senior residents may have had more motivation to complete forms due to the largely positive feedback they obtained on their O-SCOREs, while junior residents could have been discouraged by lower scores and subsequently lost their motivation to seek out additional feedback. Also, since senior residents had more experience with different orthopedic procedures, they may have felt more comfortable requesting feedback on a broader list of procedures while junior residents may have felt less comfortable requesting feedback on a procedure they did not feel comfortable performing. One resident's comment during the post-study opinion polls alluded to this phenomenon by stating "It is easy for residents to 'cherry-pick' cases that they felt went well because (1) it is always nicer to hear positive feedback than negative feedback, and (2) no one wants their program administrators to see (their) negative feedback."

Motivating residents to actively seek feedback remains a complex problem. While residents at our program are encouraged to take responsibility for their growth and maturation during residency training, we acknowledge that multiple factors may influence an individual's level of comfort with seeking direct performance feedback from an authority figure. This reluctance to seek feedback may understandably decrease compliance with a voluntary feedback program such as the one implemented herein. One solution to this problem would be to require a certain number of completed structured feedback forms for matriculation. Alternatively, programs could incentivize individual compliance through implementation of a merit-based reward system. Simple reassurance that lower scores would not trigger punitive outcomes may also be sufficient. Ultimately, the goal of immediate surgical feedback is to quantify proficiency as well as highlight areas of deficiency, so that programs may redirect resources to improve resident competence.

The technological constraints of the web-based version of the O-SCORE were a frequent complaint and limitation of the feedback program. For the sake of maintaining study participant confidentiality and for centralizing data collection, REDCap was utilized as a secure format to collect feedback data for each resident. A major limitation of the REDCap platform was that it did not offer individual access for each subject to review their O-SCORE surveys immediately following faculty completion. Due to the lack of administrative manpower to send survey results to each resident after each week, residents were only given a summary of their submitted feedback forms at the 8-week and 16-week time period.

Upon recognizing this limitation, residents and faculty were encouraged to verbally review the feedback forms at the time of completion, which occurred in approximately 78% of cases; however, the limited ability to freely review documented feedback outside of the face-to-face encounter remained a reported drawback of the web-based program as implemented in this study. This limitation may have affected motivation among subjects to seek completion of a feedback form that they would only be able to review infrequently.

Based on the operational challenges and successes of this study and the feedback provided by both faculty and residents, we would make several recommendations for orthopedic residencies that may be considering implementation of an immediate feedback program.

1. The assessment form should be concise and utilize entrustability language that offers objective feedback without use of a comparison scale of performance to peers.
2. The form should offer immediate access to completed forms for residents to review feedback at their discretion.
3. Programs should start by requesting 1 form per week to avoid compliance fatigue among participating residents and faculty. Once implemented, increasing the frequency of form submission or tailoring the frequency to the surgical volume of a given subspecialty (i.e., higher frequency among higher volume subspecialties such as sports medicine or hand) should be considered.
4. Electronic reminders to complete feedback forms should be automatic to minimize administrative burden, but should not be so frequent that they produce "alarm fatigue," as too many notifications or reminders may breed contempt among participants and reduce cooperation.
5. Flexibility to request assessment of orthopedic procedures other than ACGME Milestone procedures may encourage the habit of requesting more frequent feedback requests, which in turn may improve long-term compliance with feedback forms.

There are several strengths of this study. We confirmed the construct validity of the O-SCORE outside of its original institution. The study provides a comprehensive descriptive analysis of the positive and negative perceptions of a newly implemented surgical assessment program. Furthermore, we reported participation from nearly all residents and teaching faculty at our institution, which provides a realistic look at the challenges that may occur during the early execution phase of such a program. While several of the perceptions may be unique to the O-SCORE, they underscore common challenges that affect the implementation of surgical assessment forms in orthopedic residency programs. In our study, residents reported feeling uncomfortable asking faculty for feedback, and faculty bemoaned the often delayed feedback requests by residents. While tangible incentives (awards, scheduling prioritization) may have boosted compliance in the early phases of implementation, they may not be sustainable over the long term. Interventions to improve compliance should focus on instilling a culture that encourages seeking timely surgical feedback among residents and faculty. The best way to self-motivate residents to seek timely feedback remains a question for future studies. Nevertheless, we have been able to use the feedback provided from this study to improve the language and structure of our evaluation of surgical skills of residents at our institution.

We can conclude that implementation of the O-SCORE immediate feedback program utilizing an electronic platform is achievable and reproducible. However, technological constraints as well as compliance with both requesting and obtaining feedback among both residents and faculty physicians must temper optimism during the early stages of implementation. Programs must remain flexible and adaptable to the needs of both residents and faculty for the successful launch of a structured surgical feedback program.

ETHICAL APPROVAL

Carolinas HealthCare System IRB approved this study (IRB 12-17-28-EX, reviewed 12/29/2017).

DISCLAIMER

None.

PREVIOUS PRESENTATIONS

None.

DATA SHARING

N/A.

APPENDIX

O-SCORE
Table A1

TABLE A1. Recommended Procedures for Evaluation

Trauma	• ORIF/IMN diaphyseal tibia/femur*,†
	• ORIF ankle fracture*,†
	• Distal radius fracture*,†
	• Operative fixation hip fracture (IMN, DHS, CRPP)*,†
Pediatrics	• Supracondylar humerus fracture*,†
	• CRPP/ORIF pediatric femur fracture
	• Flexible nailing (femur, tibia, forearm)
Spine	• Decompression and lumbar interbody fusion*,†
	• ACDF
Tumor	• Prophylactic nailing of long bone*
	• Treatment of benign bone lesions (curettage, grafting, etc)
Hand	• Carpal tunnel release*,†
	• Distal radius fracture*,†
Foot and ankle	• ORIF ankle fracture*,†
	• Foot or ankle arthrodesis*,†
	• LE amputation (TMA, BKA, AKA)
Adult reconstruction	• Primary TKA*,†
	• Primary THA*,†
Sports and shoulder	• Knee arthroscopy*,†
	• Shoulder arthroscopy*,†
Other	• Other

*ACGME Milestone procedures.
†ACGME case log minimum procedure.

REFERENCES

1. Jensen AR, Wright AS, Kim S, Horvath KD, Calhoun KE. Educational feedback in the operating room: a gap between resident and faculty perceptions. *Am J Surg*. 2012;204:248–255. https://doi.org/10.1016/j.amjsurg.2011.08.019.

2. Rose JS, Waibel BH, Schenarts PJ. Disparity between resident and faculty surgeons' perceptions of preoperative preparation, intraoperative teaching, and postoperative feedback. *J Surg Educ*. 2011;68:459–464. https://doi.org/10.1016/j.jsurg.2011.04.003.

3. Rekman J, Gofton W, Dudek N, Gofton T, Hamstra SJ. Entrustability scales: outlining their usefulness for competency-based clinical assessment. *Acad Med*. 2016;91:186–190. https://doi.org/10.1097/ACM.0000000000001045.

4. Cooney CM, Cooney DS, Bello RJ, Bojovic B, Redett RJ, Lifchez SD. Comprehensive observations of resident evolution: a novel method for assessing procedure-based residency training. *Plast Reconstr Surg*. 2016;137:673–678. https://doi.org/10.1097/01.prs.0000475797.69478.0e.

5. Cate ten O. Entrustability of professional activities and competency-based training. *Med Educ*. 2005;39:1176–1177. https://doi.org/10.1111/j.1365-2929.2005.02341.x.

6. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. *Acad Med*. 2012;87:1401–1407. https://doi.org/10.1097/ACM.0b013e3182677805.

7. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381. https://doi.org/10.1016/j.jbi.2008.08.010.

8. Review committee for orthopaedic surgery—ACGME. Case log guidelines. ACGME ACFGME, ed. April 2014:1-5.

9. Cook MR, Watters JM, Barton JS, et al. A flexible postoperative debriefing process can effectively provide formative resident feedback. *J Am Coll Surg*. 2015;220:959–967. https://doi.org/10.1016/j.jamcollsurg.2014.12.048.

10. Wagner JP, Chen DC, Donahue TR, et al. Assessment of resident operative performance using a real-time mobile web system_ preparing for the milestone age. *J Surg Educ*. 2014;71:e41–e46. https://doi.org/10.1016/j.jsurg.2014.06.008.

11. Shah N, Thompson B, Averill P, Searle N. Increasing the rate of return of resident rotation evaluations by their attending physicians in an in-patient psychiatric facility. *Acad Psychiatry*. 2007;31:439–442. https://doi.org/10.1176/appi.ap.31.6.439.

12. Blum JM, Kheterpal S, Tremper KK. A comparison of anesthesiology resident and faculty electronic evaluations before and after implementation of automated electronic reminders. *J Clin Anesth*. 2006;18:264–267. https://doi.org/10.1016/j.jclinane.2005.10.008.

13. Kannappan A, Yip DT, Lodhia NA, Morton J, Lau JN. The effect of positive and negative verbal feedback on surgical skills performance and motivation. *J Surg Educ*. 2012;69:798–801. https://doi.org/10.1016/j.jsurg.2012.05.012.