Assessing Fundamental 2-Dimensional Understanding of Basic Soft Tissue Techniques

Noel Jabbour, MD, Eric J. Dobratz, MD,† Harley S. Dresner, MD,*‡ and Peter A. Hilger, MD*§*

*Department of Otolaryngology, Division of Facial Plastic and Reconstructive Surgery, University of Minnesota, Minneapolis, Minnesota; †Department of Otolaryngology Head and Neck Surgery, Division of Facial Plastic and Reconstructive Surgery, Eastern Virginia Medical School, Norfolk, Virginia; ‡Health Partners Specialty Care Center, St. Paul, Minnesota; and §Facial Plastic and Reconstructive Surgery Specialists, Edina, Minnesota

OBJECTIVE. To develop a written practical examination and scoring system for assessing trainee skills in basic soft-tissue techniques.

DESIGN. A brief written practical examination was developed to assess the ability of trainees to sketch preoperative plans and postoperative results for common soft-tissue techniques: simple-excision, M-plasty, geometric broken line closure, Z-plasty, V-to-Y flap, and rhombic flap. A scoring system was developed to assign 0 to 5 points to each of 10 items on the examination for a total score of 0–50. The 15-minute examination was administered as a pretest, posttest, and 3-month posttest assessment as part of a soft-tissue course at our institution.

SETTING. University of Minnesota, Otolaryngology Department.

RESULTS. Three raters reviewed all examination answer sheets independently. The pretest scores of examinees correlated strongly with their level of training; the average pretest for junior residents (PGY 1-2) compared with senior residents (PGY 4-5) was 17.3 (of 50) versus 26.0 ($p < 0.01$). The scoring system showed a high intrarater reliability and high interrater reliability with correlation coefficients of $r = 0.99$ and $r = 0.95$, respectively and agreement coefficients of $\kappa = 0.82$ and $\kappa = 0.77$, respectively.

CONCLUSION. This written practical examination and scoring system may be used to assess the skills of trainees accurately in basic soft tissue techniques and to expose areas of deficiency that can be addressed in future training sessions. (J Surg 68: 178-184. © 2011 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: surgical flaps, clinical competence, specialties, surgical, education, medical, graduate, OSATS, Objective Structured Assessment of Technical Skills

COMPETENCY: Medical Knowledge

INTRODUCTION

A variety of factors in the past decade has contributed to an increased interest in the development of objective measures to assess the technical skills of surgical trainees. Because of duty-hour regulations, the available hours for a surgical trainee to acquire technical skills has been concentrated, if not truncated. Indeed, in the United Kingdom, some have reported that surgical case numbers for trainees have been nearly halved.[1] Significant attention has thus been turned to the development of summative evaluations of surgical trainees for certification and recertification. Some have developed objective assessments of technical skills (OSATS) to select successful surgical trainees among applicants for subsequent surgical training.[2,3] There is also interest in the development of objective assessment tools that may be used for formative evaluations of trainees, to assess trainee progress and need for remediation.[4]

OSATS have been well described to assess trainee skills in basic soft tissue techniques, such as suturing, knot tying, excision of a lesion, performing a Z-plasty, or harvesting a skin graft.[2,5-7] Significantly less attention has been spent on assessing trainee skills in preoperative planning.

Although the technical skills, dexterity, and proper naming and use of instruments is essential to surgery, we agree with previous authors who have highlighted the importance of preoperative planning as the cornerstone to the successful execution of basic soft tissue techniques. In their description of a novel method for teaching basic skin flaps, Dinsmore and North[8] highlighted eloquently the importance of surgical planning in the execution of basic skin flaps. The manual dexterity required to complete the technical skills involved in skin flaps should be expected to improve throughout surgical training and may indeed have a ceiling effect by years 3–4 of training.[9]

Correspondence: Inquiries to Noel Jabbour, MD, Department of Otolaryngology, University of Minnesota, 420 Delaware Street SE, Mayo Mail Code 396, Minneapolis, MN 55455; fax: (612) 625-2101; e-maila; jabb00005@umn.edu

Presented at the American Academy of Facial Plastic and Reconstructive Surgery Annual Fall Meeting, October 3, 2009, San Diego, California.

University of Minnesota Soft-Tissue Course Questionnaire

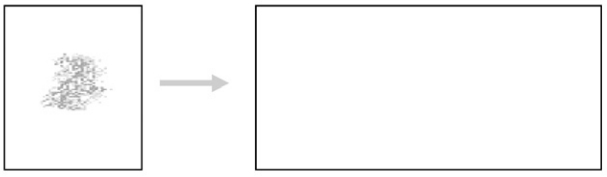
Instructions: Please complete the following exercises assuming horizontal relaxed skin-tension lines (parallel to the top and bottom of the page). For each item, please avoid drawing outside of the rectangle provided.

1. Draw a preoperative plan for the most favorable excision of this lesion shown below in order to perform primary closure.

How confident would you feel in teaching this skill to a junior resident or medical student?

1	2	3	4	5
Not at all confident		Moderately confident		Very confident

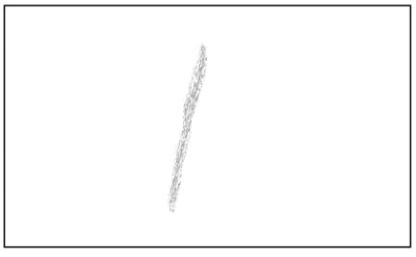
2. Draw a preoperative plan for excising the lesion shown below (left) using a M-plasty, in order to avoid extending the incision beyond the edges of the rectangle. Draw the appearance of the resulting scar in the rectangle to the right.



How confident would you feel in teaching this skill to a junior resident or medical student?

1	2	3	4	5
Not at all confident		Moderately confident		Very confident

3. Draw a preoperative plan for revising the scar shown below using Geometric Broken Line Closure:



How confident would you feel in teaching this skill to a junior resident or medical student?

1	2	3	4	5
Not at all confident		Moderately confident		Very confident

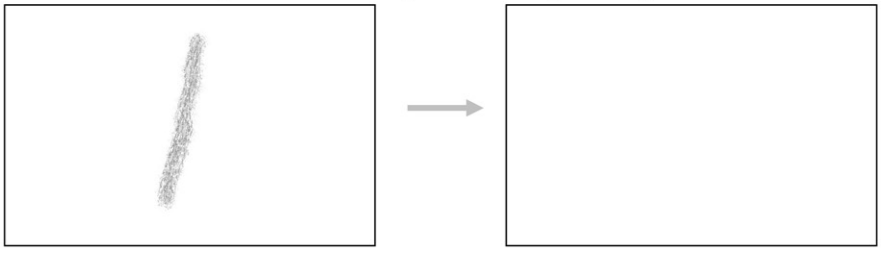
FIG. 1. Written practical examination.

However, the preoperative planning for the successful application of these basic soft tissue techniques involves complex reasoning and 3-dimensional mental reconstruction. An understanding of the basic geometry of these flaps or excision techniques is foundational to the subsequent understanding and application of these techniques. An assessment tool is needed that can focus specifically on a trainee's understanding of these techniques at this basic, 2-dimensional geometry level.

The disadvantage of currently described OSATS for basic skin flaps and scar revision techniques is that they are time consuming and are not designed to isolate preoperative planning. These examinations most often require that an expert observe the examinee either in real time or via video. In contrast, a written, practical examination would allow for an expert to rate the examination at another time and would provide subsequent standardization of the test administration. Because such an examination would require minimal time per question, many different techniques may be tested and still allow for completion in a relatively short amount of time for both the examinee and the expert rater.

Our goal is to develop a written, practical examination and scoring system that can be used to assess trainees' ability to design 2-dimensional preoperative plans and postoperative results for basic skin flaps and scar revision techniques. Such a tool has applicability for surgical trainees of all fields, especially general surgery, plastic surgery, and facial plastic surgery, as well as for primary care trainees learning to perform minor surgical procedures.

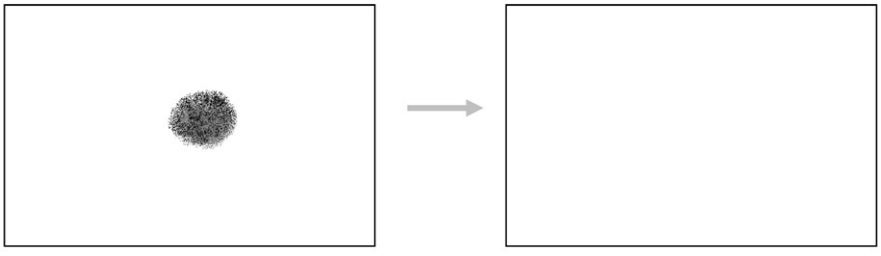
4. Draw a preoperative plan for the most favorable Z-plasty to reorient the scar shown below (left) considering relaxed skin tension lines. Then draw the appearance of the resulting scar in the rectangle to the right.



How confident would you feel in teaching this skill to a junior resident or medical student?

1	2	3	4	5
Not at all confident		Moderately confident		Very confident

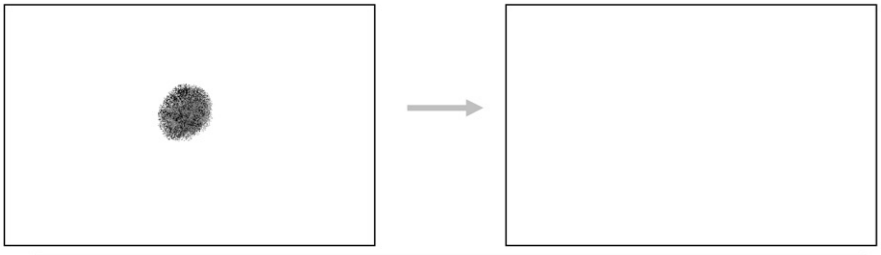
5. Draw the preoperative plans for a V-to-Y flap to close the defect below (left) considering relaxed skin tension lines. Then draw the appearance of the resulting scar in the rectangle to the right.



How confident would you feel in teaching this skill to a junior resident or medical student?

1	2	3	4	5
Not at all confident		Moderately confident		Very confident

6. Draw the preoperative plans for excision of this lesion in order to close the defect with a rhombic flap, orienting the excision considering relaxed skin tension lines and lines of maximal extensibility. Draw all four possible rhombic flaps. Next, choose one flap that will orient the vector of tension along the lines of maximal extensibility and indicate it by placing an asterisk on the flap. Finally, draw the resulting scar on the right.



How confident would you feel in teaching this skill to a junior resident or medical student?

1	2	3	4	5
Not at all confident		Moderately confident		Very confident

FIG. 1. (Continued.)

METHODS

A brief written practical examination was developed to assess the ability of trainees to sketch preoperative plans and postoperative results for common soft-tissue techniques: simple-excision of a lesion, M-plasty, geometric broken line closure, Z-plasty, V-to-Y flap, and rhombic flap. Institutional Review Board exemption was confirmed for administration of this examination in conjunction with a soft tissue course at our institution, which is offered every 3 years. The 2-page examination is shown in Fig. 1.

The 1-day course, which is offered every 3 years, consisted of a morning of lectures and an afternoon of practical laboratory time. Lecture topics included wound healing, wound repair, local anesthesia, suture selection, local flaps, and scar revision. The laboratory component was divided into 2 parts: (1) a porcine foot simulation portion to practice basic suture technique, Z-plasty, rhombic flap, geometric broken line closure, and bilobed flap, and (2) a cadeveric simulation portion to practice the repair of forehead defects, nasal defects, cheek defects, and lip defects.

Participants in the course were asked to complete a precourse examination and an identical postcourse examination, and they were later asked to complete an identical 3-month follow-up

TABLE 1. The 5-Point Scoring System

Score*	Description of Answer
0	No attempt made to answer question
1	Major deviation from ideal flap design (incorrect flap design and incorrect orientation)
2	Intermediate between 1 and 3
3	Moderate deviation from ideal flap design (incorrect flap design or incorrect orientation)
4	Intermediate between 3 and 5
5	Ideal flap design (correct flap design and correct flap orientation)

*For questions with an "a" and "b" section, incorrect orientation of the flap should only be penalized on the "a" section if the orientation is maintained for the "b" section.

examination. Each examination consisted of a separate cover sheet with the trainee's name, institution affiliation, and year of training. These cover sheets were separated from the answer sheets and matched with a unique identifier to the answer sheets to maintain anonymity.

As shown in Fig. 1, the examination consisted of 6 different tasks, 4 of which had 2 parts, for a total of 10 items. After each task, the examinee is asked, "How confident would you feel in teaching this skill to a junior resident or medical student?" with a Likert scale response from 1 (not at all confident) to 5 (very confident).

A scoring system was developed to assign 0 to 5 points to each of the 10 items on the examination for a total score of 0–50. The scoring system is shown in Table 1.

An answer key was developed by two of the raters to determine the correct appearance of a full credit response for each of the 10 items. This answer key is shown in Fig. 2. Particular clarifications were made to the scoring system before rating, as follows:

- Full credit (5 points) should be given for geometric broken line closure only if the terminal ends of the planned excision will close easily.
- Full credit (5 points) should be given for a Z-plasty only if the lesion is excised and the limbs of the flap are of adequate length.
- Full credit (5 points) may be given for the rhombic flap if drawn as shown on the key or oriented with any of the sides of the rhomboid parallel to RSTLs.

Three raters reviewed all precourse and postcourse examinations independently; the raters included 1 otolaryngology resident (N.J.), 1 facial plastic and reconstructive surgery fellow (E.J.D.), and 1 facial plastic and reconstructive surgery faculty member (H.S.D.), none of whom participated as examinees. The facial plastic surgery faculty member (H.S.D.) had not participated in the design of the examination or scoring system. Using the unique identifier on each answer sheet, raters were blinded to the name of the examinee, year in training, and whether the examination was completed precourse or postcourse. Additionally, 1 rater (N.J.) reviewed precourse and postcourse examinations a second time to evaluate intrarater reliability and also reviewed independently the 3-month postcourse examinations.

Statistical Analysis

A comparison of the results of precourse scores between junior residents (PGY-1&2) and senior residents (PGY-3&4) was performed using an independent-group T-test. A comparison of precourse, postcourse, and 3-month postcourse results was made with an independent-group T-test.

Intrarater reliability was evaluated using Pearson's correlation, linear weighted kappa, percentage of complete agreement between ratings, and percentage of agreement within 1 point between ratings. Interrater reliability was evaluated using intraclass correlation (ICC), average of linear weighted kappa coefficient, percentage of complete agreement between raters, and percentage of agreement within 1 point of the mean between raters. The average linear kappa was calculated by averaging the results of the linear weighted kappa for each of the 3 possible pairs of raters.

RESULTS

A total of 20 otolaryngology residents (18 from the University of Minnesota and 2 from the Mayo Clinic) and 1 plastic surgery fellow (from the University of Minnesota) participated in this study as examinees. In all, 16 of the 20 residents completed both precourse and postcourse examinations. Fourteen residents completed the 3-month postcourse examinations. Overall, 11 residents completed all 3 examinations. A total of 37 examination answer sheets were reviewed by all 3 raters and were reviewed a second time by 1 rater.

The average pretest score for all residents was 22.4 of 50. This score improved significantly to a posttest average of 39.1 ($p < 0.0001$). At 3 months, the average score was 39.4 (Fig. 3).

A statistically significant improvement was found in the reported confidence in the ability to teach the each skill between the precourse examination and the postcourse examination ($p < 0.05$ for all questions). When averaged together, it is clear that the reported confidence in teaching the skill parallels the objective improvement in examination scores during that same time period.

Average pretest for junior residents (PGY 1–2) compared with senior residents (PGY 4–5) was 17.3 (of 50) versus 26.0 ($p = 0.008$). After taking the course, the scores for junior residents significantly exceeded pretest scores for senior residents, 37.6 versus 26.0 ($p = 0.002$).

A total of 37 examinations was used for intrarater and interrater reliability. When the examination scores were summed for a total of 50 points, the average difference between the intrarater repeated ratings was 2.1 points out of 50. This showed a correlation coefficient of $r = 0.99$ and an agreement coefficient of $\kappa = 0.82$. When viewed by question, correlation ranged from $r = 0.87$ to $r = 1.0$ (Table 2).

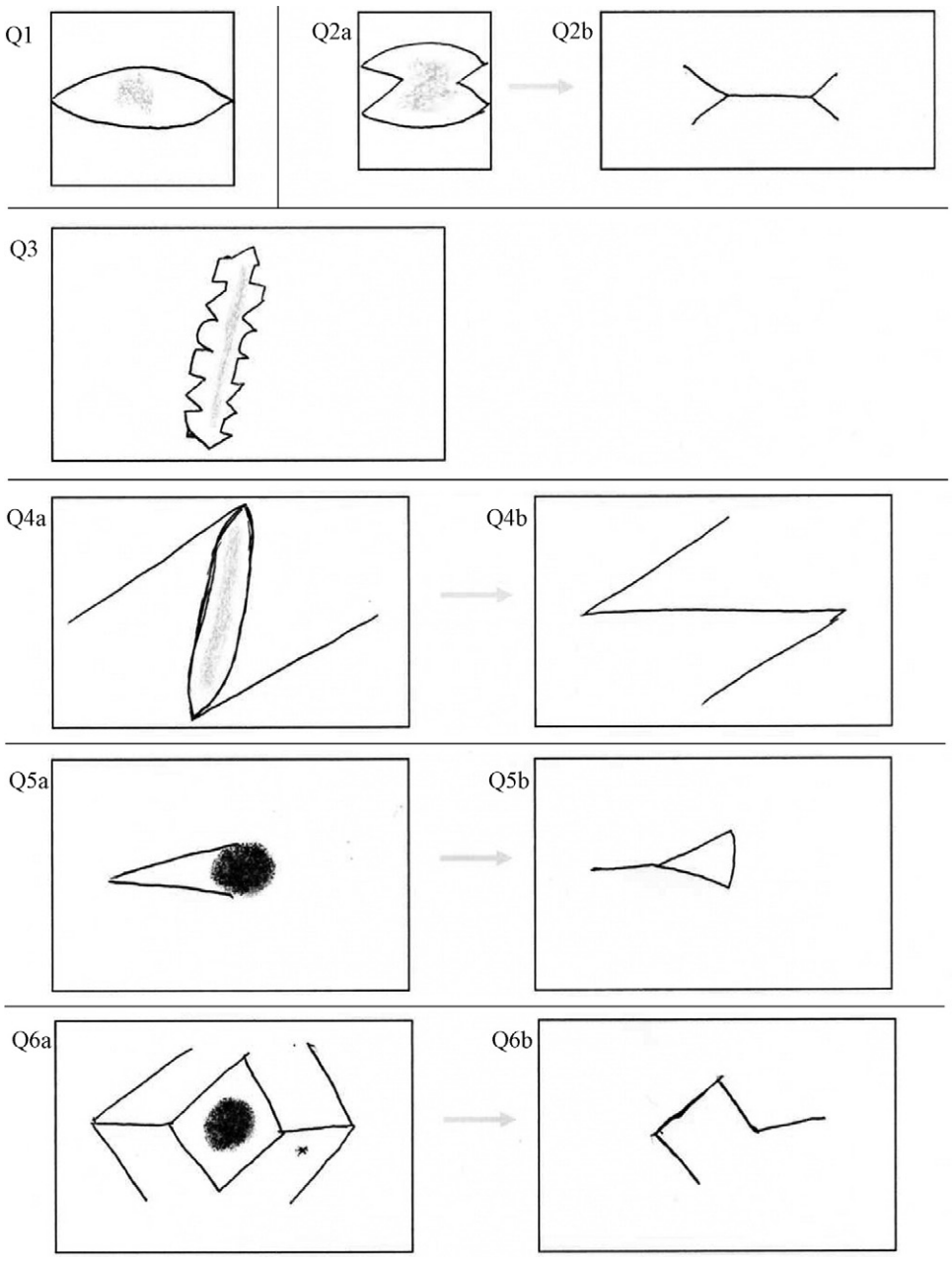


FIGURE 2. Answer key showing correct answers needed for full credit.

When examination scores are summed for a total of 50, points, the interrater reliability was high, with an ICC of $r = 0.95$ and $\kappa = 0.77$. When viewed across all questions, 87% of the time, all 3 raters either agreed on the same score from 0–5 or two raters agreed and the third was within 1 point. A breakdown of this agreement per question is shown in Table 2.

DISCUSSION

This written examination and scoring system can be used to assess objectively areas of weakness in current trainee ability. It may also be used to assess objectively what trainees have learned after an intervention, such as a course, and what they retain from this learning at a later date. Such an objective assessment tool could be used to determine which training methods are best at teaching particular skills when the interventions are given in parallel or in comparison with a control group.

Validity

In the future, decisions regarding a trainee's progress or a surgeon's certification may rest on the design of objective assessment tools and scoring systems; however, determining whether an examination and scoring system is valid is complex. A helpful review of

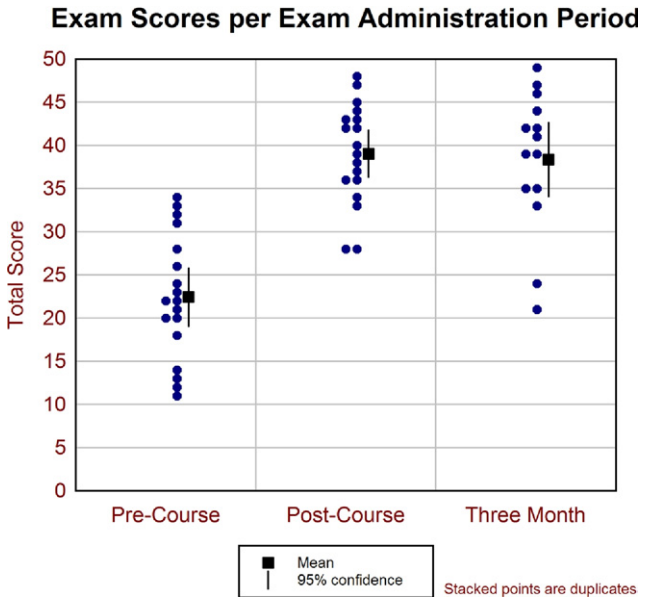


FIGURE 3. Total examination score. Distribution of scores at each examination administration period: precourse, postcourse, and 3 months postcourse.

applying rigorous scientific methods to the design of objective assessment tools in surgical education has been presented by Gallagher et al.[10] Validity involves determining that a particular test is a reliable, consistent measure of what one wishes to measure.

The purpose of this examination and scoring system is to measure objectively a trainee's ability to demonstrate and apply their understanding of the 2-dimensional geometry of basic soft tissue techniques. One would expect that senior residents would be more skilled at this than junior residents, and indeed, this examination and scoring system can detect such differences.

The reliability of the scoring system is shown in Table 2. The presentation of reliability data has the potential to be deceptive if one presents measurements of correlation, which measures the consistency of ratings, without discussing concordance, which represents agreement between ratings. Concordance measurements used for interrater reliability are unforgiving for minor disagreements between raters, such as giving a score of 4/5 versus 5/5 for a given item. This can be corrected to some degree by using linear weighting, such as Cohen's kappa as was used here. For most purposes, what one wants to know when using a Likert scoring system is how often are raters in exact agreement or in near agreement. We have chosen to present all of these measures of reliability to avoid what can be a deceptive representation if only one is presented.

The intrarater reliability overall was high, with $\kappa > 0.8$ for total examination score and > 0.8 for 8 of the 10 individual questions. The interrater reliability overall was also high with $r = 0.95$ and $\kappa = 0.77$. Agreement within 1 point occurred 87% of the time when all questions are grouped. However, from Table 2, it can be observed that the reliability can be improved upon on particular questions, such as questions 3, 4a, and 5a. When this is the case, subsequent iterations of the examination will require that reviewers should make minor modifications of the scoring system for particular items or that the instructions for these questions should be better clarified.

Limitations and Future Directions

A weakness of this study is that the same examination was used for the precourse examination, postcourse examination, and 3-month postcourse examination. Although this design allowed us to standardize the comparison of particular test items more easily, it may have enabled some trainees to demonstrate mastery of a skill on a particular test question without having mastery of the skill in general. In the future, this can be avoided if more test questions are designed for each testable skill, and these questions are used interchangeably. This will require validation

TABLE 2. Intrarater Reliability and Interrater Reliability

	IntraRater Reliability				Interrater Reliability			
Question	Pearson (r)	Kappa (κ)	Percent Exact	Percent within 1	ICC (r)	Kappa (κ)	Percent Exact	Percent within 1
Q1	0.99	0.97	97	100	0.87	0.82	73	100
Q2A	0.97	0.93	89	95	0.95	0.86	78	89
Q2B	1	0.99	97	100	0.97	0.92	92	95
Q3	0.94	0.86	81	97	0.84	0.63	68	86
Q4A	0.94	0.79	70	100	0.8	0.64	68	92
Q4B	0.94	0.81	68	95	0.91	0.77	57	89
Q5A	0.96	0.89	84	97	0.66	0.43	47	58
Q5B	0.87	0.61	51	95	0.85	0.72	92	94
Q6A	0.97	0.88	76	100	0.87	0.71	41	86
Q6B	0.96	0.86	73	95	0.88	0.76	65	78
Total score	0.99	0.82	—	—	0.95	0.77	—	—

ICC, intraclass correlation coefficient.

Percent Exact = Percent of absolute agreement between raters. Percent within 1 = Percent of evaluations for which the raters' scores differed by no more than 1 point.

of each of these new items and determination of the relative difficulty of each of the interchangeable items.

Future directions may also include application of this scoring system to more difficult written questions, such as designing local flaps and scar revision techniques that can be drawn on standardized preoperative photographs. This scoring system may also be used as a component of other assessment tools; that is, it may be used to assess the preoperative plans for basic soft tissue techniques that a trainee then performs as part of a bench examination or an OSATS.

Thus, although the examination presented here may be used "as is," it is not intended solely for this purpose. It is offered here as an example of what can be done in a written, practical examination to objectively assess trainee skills in designing basic flaps and scar revision techniques.

CONCLUSIONS

The ability to design and execute basic skin flaps and scar revision techniques is essential to surgeons of all specialties. A thorough understanding of the geometry of these flaps is needed before one can apply these skills successfully to address the needs of a particular patient.

As a written examination, this objective assessment tool is far more efficient to administer and to grade than bench examinations and has the advantage of being able to isolate preoperative planning and flap design skills. The scoring system has been shown to have high intrarater and interrater reliability.

This written, practical examination and scoring system can be used to assess objectively the skills of trainees in basic soft tissue techniques and can be used to expose areas of deficiency that can be addressed in future training sessions.

ACKNOWLEDGMENTS

The authors wish to thank Bevan Yueh, MD, MPH, Department of Otolaryngology—Head & Neck Surgery, University of Minnesota, and Connie Schmitz, PhD, Department of Surgery, University of Minnesota, for critical review of the manuscript prior to submission.

REFERENCES

1. Jaffer A, Bednarz B, Challacombe B, Sriprasad S. The assessment of surgical competency in the UK. *Int J Surg*. 2009;7:12-15.
2. Carroll SM, Kennedy AM, Traynor O, Gallagher AG. Objective assessment of surgical performance and its impact on a national selection programme of candidates for higher surgical training in plastic surgery. *J Plast Reconstr Aesthet Surg*. 2008;62:1543-1549.
3. Gallagher AG, Neary P, Gillen P, et al. Novel method for assessment and selection of trainees for higher surgical training in general surgery. *ANZ J Surg*. 2008;78:282-290.
4. Brown DJ, Thompson RE, Bhatti NI. Assessment of operative competency in otolaryngology residency: Survey of US program directors. *Laryngoscope*. 2008;118:1761-1764.
5. Collins AM, Ridgway PF, Hassan MS, et al. Surgical instruction for general practitioners: How, who and how often? *J Plast Reconstr Aesthet Surg*. 2009;63:1156-1162.
6. Khan MS, Bann SD, Darzi AW, Butler PE. Assessing surgical skill using bench station models. *Plast Reconstr Surg*. 2007;120:793-800.
7. Mackay S, Datta V, Chang A, et al. Multiple objective measures of skill (MOMS): a new approach to the assessment of technical ability in surgical trainees. *Ann Surg*. 2003;238:291-300.
8. Dinsmore RC, North JH. Basic skin flaps for the general surgeon: a teaching method. *South Med J*. 2000;93:783-786.
9. Munz Y, Moorthy K, Bann S, et al. Ceiling effect in technical skills of surgical residents. *Am J Surg*. 2004;188:294-300.
10. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc*. 2003;17:1525-1529.