Title

First-Year Results of the American Board of Anesthesiology's Objective Structured Clinical Examination for Initial Certification.

Journal

Anesthesia & Analgesia, 131(5)

ISSN

0003-2999

Authors

Warner, David O
Lien, Cynthia A
Wang, Ting
et al.

Publication Date

2020-11-01

DOI

10.1213/ane.0000000000005086

Peer reviewed

SPECIAL ARTICLE

First-Year Results of the American Board of Anesthesiology's Objective Structured Clinical Examination for Initial Certification

David O. Warner, MD,* Cynthia A. Lien, MD,† Ting Wang, PhD,‡ Yan Zhou, PhD,‡ Robert S. Isaak, DO,§ Cathleen Peterson-Layne, PhD, MD,‖ Ann E. Harman, PhD,‡ Alex Macario, MD, MBA,¶ Robert R. Gaiser, MD,# Santhanam Suresh, MD, MBA,** Deborah J. Culley, MD,†† James P. Rathmell, MD,†† Mark T. Keegan, MB, BCh,* Daniel J. Cole, MD,‡‡ Brenda G. Fahy, MD, MCCM,§§ Rupa J. Dainer, MD,‖‖ and Huaping Sun, PhD‡

In 2018, the American Board of Anesthesiology (ABA) became the first US medical specialty certifying board to incorporate an Objective Structured Clinical Examination (OSCE) into its initial certification examination system. Previously, the ABA's staged examination system consisted of 2 written examinations (the BASIC and ADVANCED examinations) and the Standardized Oral Examination (SOE). The OSCE and the existing SOE are now 2 separate components of the APPLIED Examination. This report presents the results of the first-year OSCE administration. A total of 1410 candidates took both the OSCE and the SOE in 2018. Candidate performance approximated a normal distribution for both the OSCE and the SOE, and was not associated with the timing of the examination, including day of the week, morning versus afternoon session, and order of the OSCE and the SOE. Practice-based Learning and Improvement was the most difficult station, while Application of Ultrasonography was the least difficult. The correlation coefficient between SOE and OSCE scores was 0.35 ([95% confidence interval {CI}, 0.30–0.39]; $P < .001$). Scores for the written ADVANCED Examination were modestly correlated with scores for the SOE ($r = 0.29$ [95% CI, 0.25–0.34]; $P < .001$) and the OSCE ($r = 0.15$ [95% CI, 0.10–0.20]; $P < .001$). Most of the candidates who failed the SOE passed the OSCE, and most of the candidates who failed the OSCE passed the SOE. Of the 1410 candidates, 77 (5.5%) failed the OSCE, 155 (11.0%) failed the SOE, and 25 (1.8%) failed both. Thus, 207 (14.7%) failed at least 1 component of the APPLIED Examination. Adding an OSCE to a board certification examination system is feasible. Preliminary evidence indicates that the OSCE measures aspects of candidate abilities distinct from those measured by other examinations used for initial board certification. (Anesth Analg XXX;XXX:00–00)

GLOSSARY

ABA = American Board of Anesthesiology; **CI** = confidence interval; **OSCE** = Objective Structured Clinical Examination; **SD** = standard deviation; **SOE** = Standardized Oral Examination

The American Board of Anesthesiology (ABA, Raleigh, NC) requires that candidates for initial certification in anesthesiology pass a series of examinations, consisting of the BASIC, ADVANCED, and APPLIED examinations. Both the BASIC (taken by residents at the end of the first year of clinical anesthesia training) and the ADVANCED (taken after the completion of residency) are written multiple-choice question examinations. Before 2018, the APPLIED Examination consisted of a single oral examination format, referred to as the Standardized Oral Examination (SOE), which includes case-based

From the *Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, Minnesota; †Department of Anesthesiology, Medical College of Wisconsin, Milwaukee, Wisconsin; ‡The American Board of Anesthesiology, Raleigh, North Carolina; §Department of Anesthesiology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; ‖Department of Anesthesiology, Emory University, Atlanta, Georgia; ¶Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University, Stanford, California; #Department of Anesthesiology, University of Kentucky, Lexington, Kentucky; **Department of Pediatric Anesthesiology, Ann & Robert H. Lurie Children's Hospital of Chicago, Northwestern University, Chicago, Illinois; ††Department of Anesthesiology, Perioperative and Pain Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; ‡‡Department of Anesthesiology and Perioperative Medicine, University of California, Los Angeles, Los Angeles, California; §§Department of Anesthesiology, University of Florida, Gainesville, Florida; and ‖‖Department of Ambulatory Surgery, Pediatric Specialists of Virginia, Fairfax, Virginia.

DOI: 10.1213/ANE.0000000000005086

Accepted for publication June 19, 2020.

Funding: Institutional and/or departmental.

Conflicts of Interest: See Disclosures at the end of the article.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.anesthesia-analgesia.org).

Reprints will not be available from the authors.

Address correspondence to Huaping Sun, PhD, The American Board of Anesthesiology, 4208 Six Forks Rd, Suite 1500, Raleigh, NC 27609. Address e-mail to huaping.sun@theaba.org.

patient care discussions designed to evaluate the general and subspecialty competencies required of a consultant anesthesiologist. In 2018, the ABA expanded the APPLIED Examination by adding an Objective Structured Clinical Examination (OSCE) component.

With the administration of the first OSCE in March 2018, the ABA became the first medical specialty certifying board in the United States to incorporate this type of assessment into its examination system. Miller's Pyramid of Assessment provides a framework to assess clinical competencies, defining progressive stages of "Knows (Knowledge)," "Knows how (Competence)," "Shows how (Performance)," and "Does (Action)."[1] The written BASIC and ADVANCED examinations aim to assess medical knowledge at the foundational level of "Knows," and the SOE aims to assess application of medical knowledge and other more complex skills and abilities such as clinical reasoning and judgment at the level of "Knows how." The primary rationale for adding an OSCE component was to allow the candidate to "show how" they might actually "do" in a simulated practice setting.[1,2] Inherent in this rationale is that the OSCE would capture competencies expected of an ABA diplomate different from those assessed in the ABA's written examinations or the SOE.

The format, administration, and scoring of the ABA OSCE have been described previously.[2] The purpose of this report is to present the results of the first year of the OSCE administration, including candidate performance and pass rates, the association between OSCE administration logistics and candidate performance, and the relative difficulty of individual OSCE stations. Performances by candidates on the OSCE, the SOE, and ADVANCED examinations were also compared to provide preliminary evidence as to whether these 3 examinations measure different aspects of candidate abilities.

METHODS

This study was determined to be exempt from review by the Mayo Clinic Institutional Review Board (Rochester, MN).

Description of the SOE and the OSCE

The administration and scoring of both the SOE and OSCE were previously described.[2,3] For the SOE, 3 modules are assessed in each of the two 35-minute examination sessions by means of case-based discussions of pre-, intra-, and postoperative anesthesiology care and nonoperative anesthesiology subspecialty care. For the OSCE, the tasks consist of a circuit of 7 stations that assess Communication and Professionalism Skills (5 stations selected each week from 7 available stations according to an examination blueprint) and Technical Skills (2 stations selected from 3 available stations) as defined by the OSCE Content Outline (Supplemental Digital Content, Material 1, http://links.lww.com/AA/D160).

Global ratings are used for each module in the SOE and each station in the OSCE on a 4-level ordinal scale—whether a candidate "consistently," "often," "occasionally," or "rarely" demonstrates the characteristics expected of an ABA diplomate. For the SOE, 2 examiners in each examination session independently rate the candidates on each of the 3 modules, which results in a total of 12 ratings from 4 examiners. For the OSCE, each examiner typically focuses on 1 station, with multiple examiners rating the same station in 1 examination week. For a given candidate, their performance on each station is rated by a different examiner, for a total of 7 ratings from 7 examiners. For candidates whose initial total ratings are in the lower range of the distribution, Communication and Professionalism stations are double scored, resulting in a total of 12 ratings from 12 examiners—10 ratings for 5 Communication and Professionalism stations and 2 ratings for 2 Technical Skills stations. The technical stations are not double scored because expected behaviors and answers are more objective and less dependent on examiner judgment.[2] The ratings are used to estimate candidate ability using the many-facet Rasch model.[2–4] This model provides estimates of candidate ability, how difficult an examiner is in their grading (ie, examiner severity), and task/station difficulty, which can be expressed as scaled scores. Higher scores represent more able candidates, more severe examiners, and more difficult tasks/stations. Examiner severity is estimated separately for SOE and OSCE ratings.

The ABA sets a criterion-based passing standard for each examination based on standard-setting procedures as previously described.[2,3] This standard requires the passing candidate to on average "often" demonstrate expected diplomate characteristics, including an adjustment factor based on the standard error of measurement to account for potential measurement error. Once the standards are set (separately for the SOE and the OSCE), they are applied to all subsequent examinations. For each examination week, separate Rasch models for the SOE and the OSCE were run based on that week's examination data, allowing for timely independent pass/fail decisions for each component of the APPLIED Examination.

Statistical Analyses

This report included data from all candidates who took the APPLIED Examination for the first time in 2018 during 1 of 9 separate 4-day examination weeks. Each candidate contributed 1 ADVANCED score, 1 SOE score, and 1 OSCE score to the analysis. All candidates took both components (the SOE and the OSCE) either

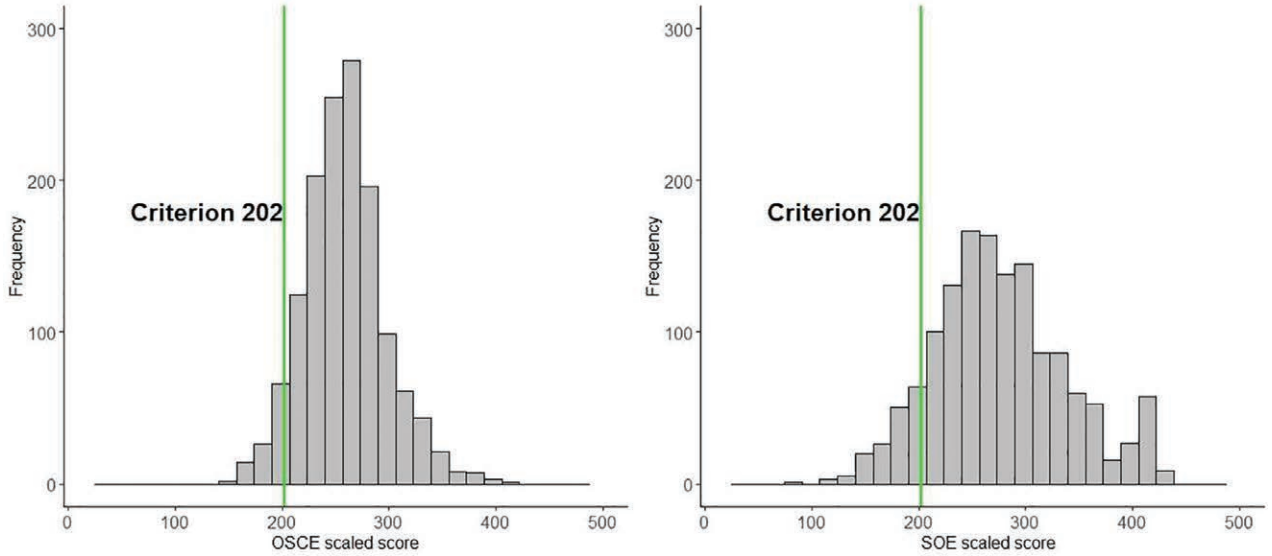

Figure 1. Histogram of the distribution of scaled scores for the OSCE (A) and the SOE (B). Vertical lines denote the scaled scores of the passing standard (ie, the passing criterion). OSCE indicates Objective Structured Clinical Examination; SOE, Standardized Oral Examination.

in a morning session or an afternoon session, with half taking their OSCE first and the other half taking their SOE first during each session. Separate descriptive analyses were conducted to evaluate the association between examination administration logistics and candidate performance, including the day of the week and the time of the examination session (morning or afternoon), and the order examinations were given (the OSCE before the SOE or vice versa).

To rank the relative difficulty of each OSCE station, difficulties for each station were estimated in a single Rasch model incorporating all data from the 9 examination weeks. For those candidates who failed the OSCE, their average scaled score on each station was calculated to identify whether there were stations that posed particular difficulty for failing candidates.

Correlation between candidates' SOE and OSCE scaled scores was evaluated by calculating the Pearson product-moment correlation coefficient. Correlations were also calculated between written ADVANCED Examination scaled scores and the OSCE or the SOE scaled scores. For candidates who required more than 1 attempt to pass the ADVANCED Examination, scores on their first attempt were utilized.

SOE and OSCE scores were obtained from Facets Rasch Software 3.71.4 (Beaverton, OR). Other statistical analyses were performed using R 3.5.3 (R Foundation for Statistical Computing, Vienna, Austria). *P* values were based on 2-tailed statistical testing and a *P* value of <.05 was considered to indicate statistical significance.

RESULTS

A total of 1410 candidates completed both the SOE and the OSCE in 2018. For the OSCE, the distribution of candidate abilities estimated from Rasch models approximated a normal distribution. The mean scaled score was 259 with a standard deviation (SD) of 38 (Figure 1). For the SOE, the distribution of candidate abilities estimated from Rasch models also approximated a normal distribution, with a mean of 276 and an SD of 63 (Figure 1).

Table 1. Candidate Performance According to Examination Timing for OSCE and SOE

	OSCE	SOE	N
Examination day			
1	258 (38)	276 (59)	458
2	258 (37)	273 (62)	449
3	262 (38)	277 (66)	407
4	255 (45)	293 (62)	96
Examination session			
Morning	257 (39)	276 (65)	727
Afternoon	261 (37)	276 (60)	683
Examination order			
OSCE first	259 (38)	279 (64)	700
SOE first	258 (38)	274 (61)	710

Values are expressed as scaled scores, mean (standard deviation). Statistical comparisons were not performed because the data are not sampled, but represent the entire population.

Abbreviations: OSCE, Objective Structured Clinical Examination; SOE, Standardized Oral Examination.

There was no evidence that candidate performance was associated with the timing of the examination, including day of the week, morning versus afternoon administration, and order of OSCE and SOE because there was little variation in mean candidate scores (Table 1).

Using a single Rasch model based on all data from the 9 examination weeks, the relative difficulty of each OSCE station was estimated (Table 2). Among the 7 Communication and Professionalism stations, the most difficult was Practice-based Learning and Improvement

Table 2. Difficulty of OSCE Stations and Performance of Failing Candidates

Station	Station Difficulty as Scaled Score (Mean [95% CI], N)[a]	Failing Candidates' Scaled Score (N = 77, Mean [SD])[b]
Communication and Professionalism		
Practice-based Learning and Improvement	193 (190–196), N = 1430	184 (50)
Treatment Options	182 (179–185), N = 1487	177 (49)
Periprocedural Complications	178 (175–181), N = 2154	192 (49)
Informed Consent	169 (166–172), N = 2133	183 (56)
Ethical Issues	169 (165–173), N = 1384	173 (63)
Communication with Other Professionals	153 (150–157), N = 2154	194 (41)
Technical Skills		
Interpretation of Monitors	189 (184–193), N = 757	222 (49)
Interpretation of Echocardiograms	153 (148–158), N = 654	201 (62)
Application of Ultrasonography	147 (143–152), N = 1410	189 (55)

Abbreviations: CI, confidence interval; OSCE, Objective Structured Clinical Examination; SD, standard deviation.

[a]The relative difficulty of each station was estimated using a single Rasch model based on data from all 9 examination weeks. Higher scaled scores indicate greater difficulty. N indicates the number of ratings for each station in the Rasch model.

[b]Values for each station were calculated based on 77 candidates who failed the OSCE. For reference, the passing criterion for the overall OSCE was a scaled score of 202.

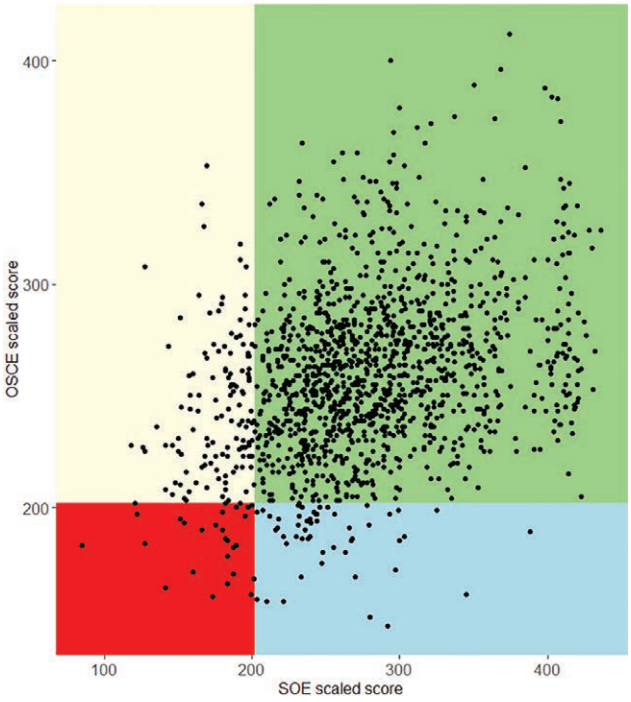

Figure 2. Relationship between SOE and OSCE scaled scores. Four quadrants depict pass-fail decisions based on the passing standard of 202 for both the SOE and the OSCE. Green: Passed both the SOE and the OSCE; red: failed both the SOE and the OSCE; yellow: failed the SOE but passed the OSCE; blue, passed the SOE but failed the OSCE. OSCE indicates Objective Structured Clinical Examination; SOE, Standardized Oral Examination.

(a scaled score of 193), and the least difficult was Communication with Other Professionals (a scaled score of 153). Among the 3 Technical Skills stations, Interpretation of Monitors was the most difficult (a scaled score of 189), and Application of Ultrasonography was the least difficult (a scaled score of 147).

For candidates who failed the OSCE (N = 77), among the Communication and Professionalism stations, their mean scaled score was the highest for Communication with Other Professionals (indicating better performance)

Table 3. Pass/Fail Results for the OSCE and the SOE in 2018

	OSCE Fail (%)	OSCE Pass (%)	Total (%)
SOE fail	25 (1.8)	130 (9.2)	155 (11.0)
SOE pass	52 (3.7)	1203 (85.3)	1255 (89.0)
Total	77 (5.5)	1333 (94.5)	1410

Percentages are calculated based on the total number of candidates (N = 1410).

Abbreviations: OSCE, Objective Structured Clinical Examination; SOE, Standardized Oral Examination.

and the lowest for Ethical Issues (Table 2). Among the Technical Skills stations, their mean scaled score was the highest for Interpretation of Monitors and the lowest for Application of Ultrasonography (Table 2).

The scatterplot between SOE and OSCE scores (Figure 2) demonstrated a significant correlation between these 2 components (r = 0.35 [95% confidence interval {CI}, 0.30–0.39], $P < .001$, r^2 = 0.12). However, for a given level of SOE performance, there was considerable variability in candidates' OSCE performance, and vice versa; only 12% of the variability in the scores for 1 examination was accounted for by the scores in the other examination. The passing standard in 2018 for both the SOE and the OSCE was set at a scaled score of 202 (used to generate the shaded quadrants depicted in Figure 2). Most of the candidates who failed the SOE passed the OSCE (upper left quadrant of Figure 2), and most of the candidates who failed the OSCE passed the SOE (lower right quadrant of Figure 2). Some candidates failed both the SOE and the OSCE (lower left quadrant of Figure 2). Of the 1410 candidates examined in 2018, 77 (5.5%) failed the OSCE, 155 (11.0%) failed the SOE, and 25 (1.8%) failed both (Table 3). Thus, 207 (14.7%) failed at least 1 component of the APPLIED Examination (ie, the OSCE and/or the SOE), and are required to retake and pass the component they failed to achieve initial board certification.

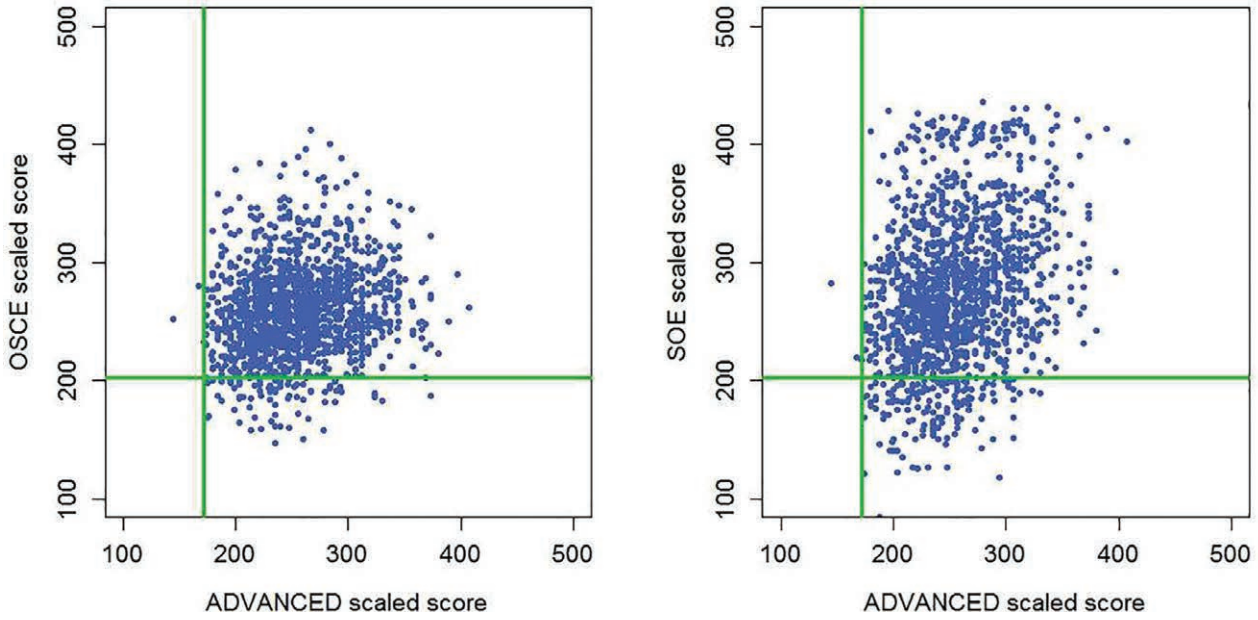

Figure 3. Relationship between the ADVANCED (written) Examination and OSCE (A) and SOE (B) scaled scores, respectively. Lines indicate the scaled scores of the passing standard (ie, the passing criterion). For example, the points below the horizontal lines represent the candidates failing the OSCE and the SOE, respectively. Candidates must pass the ADVANCED Examination to be eligible to take the APPLIED Examination. The few data points to the left of the vertical line reflect those candidates who initially failed, but subsequently passed the ADVANCED Examination—their first-attempt ADVANCED scaled scores are presented. OSCE indicates Objective Structured Clinical Examination; SOE, Standardized Oral Examination.

There was a weak, but statistically significant, correlation between ADVANCED Examination (ie, written examination) and OSCE scores (r = 0.15 [95% CI, 0.10–0.20], P < .001, r^2 = 0.02, Figure 3), implying that for a given level of performance on the ADVANCED Examination, there was considerable variability in candidates' OSCE performance and vice versa. The correlation between the ADVANCED Examination and the SOE scores was also relatively weak but statistically significant (r = 0.29 [95% CI, 0.25–0.34], P < .001, r^2 = 0.09, Figure 3). Performance on the ADVANCED Examination accounts for approximately 2% and 9% of the OSCE and the SOE variability, respectively.

DISCUSSION

In 2018, the ABA became the first medical specialty certifying board in the United States to include an OSCE in its certification examination system. This report presents the results from the first year of its administration.

Scoring of both the SOE and the OSCE uses global ratings given by examiners based on their assessment of how candidates perform on specific tasks in the SOE and individual stations in the OSCE. The rationale for the choice of this method has been previously discussed.[2–4] Since 2002, ratings for the SOE have been analyzed using a model that provides estimates of candidate ability, controlling for variations in task difficulty and examiner severity.[3,4] The same analytic method was chosen for the OSCE.[2] The distribution of candidate ability expressed in scaled scores approximated a normal distribution for both the SOE and the OSCE, with some clustering of scores at the higher end of the distribution for the SOE. This clustering likely reflects candidates receiving the highest ratings of "frequently" for all SOE tasks and does not impact pass/fail decisions because these are clearly high-performing candidates.

One concern that arose during the design of the OSCE was whether candidate performance was associated with the timing of taking the components of the APPLIED Examination. For example, if candidates found the OSCE to be particularly stressful, their performance on the subsequent SOE component could be impaired if they took the OSCE first (or vice versa). However, there was no evidence that candidate performance was associated with the timing of APPLIED examination components (Table 1). With the reasonable assumption that the actual distribution of candidate abilities does not depend on these timing factors, this finding suggests that candidates receive a comparably fair examination experience regardless of the timing of their examination.

Individual items vary in difficulty in any examination. This was applicable to individual stations within the OSCE. Although these differences were relatively modest, it was interesting to note that Practice-based Learning and Improvement was the most difficult of the 7 Communication and Professionalism skills (ie, station difficulty expressed as a scaled score was relatively high). It was somewhat surprising that Interpretation of Monitors, a skill required for the daily practice of anesthesiologists, also challenged

candidates. These findings are potentially valuable both in the design of future OSCE scenarios and as feedback to training programs to guide residency curriculum. On the other hand, it was perhaps heartening to note that candidates were rated relatively highly on Communication with Other Professionals (ie, station difficulty was relatively low), a crucial skill for practicing anesthesiologists. Candidates were on average rated as relatively proficient in the Interpretation of Echocardiograms and Application of Ultrasonography, both of which are more recent additions to the residency curriculum.

The ABA incorporates multiple examinations as part of the requirements for initial certification, each designed to measure distinct candidate abilities that are important to anesthesiology practice. Prior studies provide evidence that each certifying examination indeed measures distinct abilities. For example, the results of both the written examination and the SOE are independently correlated with ratings of resident clinical performance by faculty during their final year of residency.[5] Another study found that candidates who passed the SOE had a lower risk of subsequent prejudicial actions against their medical licenses; no such association was found for passing the written examination.[6] Part of the rationale for introducing the OSCE is the expectation that it would assess aspects of candidate abilities distinct from those assessed in existing examinations.[2] The relationship between the scores of individuals on different examinations can provide some insight into whether these examinations measure distinct candidate abilities. When candidate score pairs are graphically depicted for any 2 examinations (Figures 2 and 3), the observed variations from a linear relationship suggest that the examinations measure different abilities, that there is measurement error associated with each examination (eg, individual scores do not perfectly reflect the ability of a given candidate), or a combination of both factors.

There was only a weak correlation between candidate scores on the ADVANCED Examination and both the OSCE and the SOE (Figure 3). Thus, some individuals who scored quite well on the written ADVANCED Examination scored poorly on the OSCE and the SOE, and vice versa; that is, performance on the ADVANCED Examination may not predict performance in the OSCE or SOE. These results are consistent with a prior report that analyzed the relationship between scores on the SOE and the written examination taken at the conclusion of residency training (correlation coefficient of 0.31 based on data from 2012 to 2017).[3] Recognizing that these correlations are truncated because of the requirement that candidates must pass the ADVANCED Examination to attempt the OSCE and the SOE, this finding provides preliminary evidence that each component of the ABA APPLIED Examination assesses different aspects of candidate abilities compared with the written ADVANCED Examination.

Performance on both the SOE and the OSCE depends on the candidate's level of clinical knowledge and their ability to communicate, so it is not surprising that SOE and OSCE scores are correlated. When candidate SOE and OSCE score pairs are graphically depicted, there was considerable variation from a linear relationship between individual candidate scores on these 2 examinations (Figure 2). The result that the majority of candidates who failed 1 component of the APPLIED Examination passed the other (Figure 2) provided additional support that the SOE and the OSCE may indeed assess different aspects of candidate abilities. Nevertheless, these results are limited to descriptive statistics; further analyses with more advanced study methodology (eg, exploratory and confirmatory factor analyses) would be necessary to formally evaluate the hypothesis that the SOE and the OSCE indeed measure different constructs related to candidate abilities.

In conclusion, this analysis presents the initial results of the first OSCE used as a part of the initial certification process for a US medical specialty certifying board, showing that this examination format is feasible to administer in a high-stakes, high-volume context. Preliminary evidence suggests that the OSCE measures constructs distinct from those measured by other examinations used in the initial certification process. Further study will be necessary to determine if there is a relationship between OSCE scores and measures of physician performance in clinical practice—recognizing that such performance may be challenging to assess on a large scale.

DISCLOSURES

Name: David O. Warner, MD.
Contribution: This author helped conceptualize the manuscript, manage, analyze, and interpret the data, and draft the manuscript.
Conflicts of Interest: D. O. Warner serves as a Director for the American Board of Anesthesiology.
Name: Cynthia A. Lien, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: C. A. Lien is a former Director of the American Board of Anesthesiology.
Name: Ting Wang, PhD.
Contribution: This author helped conceptualize the manuscript, manage, analyze, and interpret the data, and draft the manuscript.
Conflicts of Interest: T. Wang is a former staff member of the American Board of Anesthesiology.
Name: Yan Zhou, PhD.
Contribution: This author helped conceptualize the manuscript, manage, analyze, and interpret the data, and draft the manuscript.

Conflicts of Interest: Y. Zhou is a former staff member of the American Board of Anesthesiology.
Name: Robert S. Isaak, DO.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: R. S. Isaak serves as the Chair for the American Board of Anesthesiology Objective Structured Clinical Examination Committee.
Name: Cathleen Peterson-Layne, PhD, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: C. Peterson-Layne serves as the Vice Chair for the American Board of Anesthesiology Objective Structured Clinical Examination Committee.
Name: Ann E. Harman, PhD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: A. E. Harman is a staff member of the American Board of Anesthesiology.
Name: Alex Macario, MD, MBA.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: A. Macario serves as a Director for the American Board of Anesthesiology.
Name: Robert R. Gaiser, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: R. R. Gaiser serves as a Director for the American Board of Anesthesiology.
Name: Santhanam Suresh, MD, MBA.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: S. Suresh serves as a Director for the American Board of Anesthesiology.
Name: Deborah J. Culley, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: D. J. Culley serves as a Director for the American Board of Anesthesiology.
Name: James P. Rathmell, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: J. P. Rathmell serves as a Director for the American Board of Anesthesiology.
Name: Mark T. Keegan, MB, BCh.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: M. T. Keegan serves as a Director for the American Board of Anesthesiology.
Name: Daniel J. Cole, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: D. J. Cole serves as a Director for the American Board of Anesthesiology.
Name: Brenda G. Fahy, MD, MCCM.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: B. G. Fahy serves as a Director for the American Board of Anesthesiology.
Name: Rupa J. Dainer, MD.
Contribution: This author helped conceptualize the manuscript, interpret the data, and draft the manuscript.
Conflicts of Interest: R. J. Dainer serves as a Director for the American Board of Anesthesiology.
Name: Huaping Sun, PhD.
Contribution: This author helped conceptualize the manuscript, manage, analyze, and interpret the data, and draft the manuscript.
Conflicts of Interest: H. Sun is a staff member of the American Board of Anesthesiology.
This manuscript was handled by: Edward C. Nemergut, MD.

REFERENCES

1. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65:S63–S67.
2. Warner DO, Isaak RS, Peterson-Layne C, et al. Development of an objective structured clinical examination as a component of assessment for initial board certification in anesthesiology. *Anesth Analg*. 2020;130:258–264.
3. Sun H, Warner DO, Patterson AJ, et al. The American Board of Anesthesiology's standardized oral examination for initial board certification. *Anesth Analg*. 2019;129:1394–1400.
4. Linacre JM. *Many-Facet Rasch Measurement*. 2nd ed. Chicago, IL: MESA Press; 1994.
5. Baker K, Sun H, Harman A, Poon KT, Rathmell JP. Clinical performance scores are independently associated with the American Board of Anesthesiology certification examination scores. *Anesth Analg*. 2016;122:1992–1999.
6. Zhou Y, Sun H, Culley DJ, Young A, Harman AE, Warner DO. Effectiveness of written and oral specialty certification examinations to predict actions against the medical licenses of anesthesiologists. *Anesthesiology*. 2017;126:1171–1179.