REPORT ON A PRELIMINARY EVALUATION OF THE RECONFIGURED SELF-EVALUATION EXAMINATION (SEE)

Posted October 17, 2017

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

Introduction

On September 1, 2016, the NBCRNA launched a reconfigured version of the Self-Evaluation Examination (SEE). This "new" SEE featured the following distinctions from the older version of the exam ("old" SEE):

- Alignment of SEE content outline with four primary domains of NCE outline
  - Basic Sciences
  - Equipment, Instrumentation & Technology
  - Basic Principles of Anesthesia
  - Advanced Principles of Anesthesia
- Increased test length from 160 to 240 questions, to improve reliability and usefulness of domain-level scores
- Increased time limit from three hours to four hours
- Intended predictive capability with respect to future performance on the National Certification Examination (NCE)

The purpose of this report is to use accumulated data to conduct a preliminary evaluation of the new SEE with respect to the stated goals of the program.

Score and Timing Summaries

From the launch of the new SEE, through the end of May 2017, the new SEE has been administered to 2,752 examinees. This testing volume represents a 28% increase compared to the same timeframe the previous year. Table 1 contains the average total scores for the old SEE (for the entire FY 2016) and the new SEE (for FYTD 2017). The average scores are comparable[1].

Score Performance

Table 1. Comparison of Average Scores, New SEE vs. Old SEE, by Year in Program

Year in Program	**New SEE Average (SD)**	**Old SEE Average (SD)**
Year 1	397.0 (38.6)	389.5 (44.1)
Year 2	406.1 (43.0)	401.7 (42.6)
Year 3 and Up	410.2 (44.2)	398.7 (44.5)
Total	**406.5 (43.1)**	**399.3 (43.5)**

[1] While old SEE vs. new SEE score comparisons are not feasible at the *domain level* (due to the different domain specifications), it is possible to compare *overall* scores before and after the configuration changes. Both "before" and "after" scores are assumed to be aggregate measures of general anesthesia knowledge, and are derived from the same underlying logit scale.

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

Table 2 contains descriptive summaries for the overall score and domain-level scores, by year in program. More detailed information, including percentile transformations, is available in the FY2017 SEE Interpretive Guides on the NBCRNA website.

Table 2. New SEE Score Summaries, by Year in Program

	Year in Program			
Domain	**Year 1**	**Year 2**	**Year 3 and Up**	**Total**
Basic Sciences	403.0 (43.3)	401.2 (47.5)	405.2 (49.4)	402.8 (47.7)
Equipment, Instrumentation & Technology	400.9 (45.2)	405.7 (48.1)	409.1 (47.2)	406.3 (47.5)
Basic Principles of Anesthesia	398.7 (45.9)	410.5 (50.1)	413.8 (51.7)	410.3 (50.4)
Advanced Principles of Anesthesia	389.4 (46.0)	410.4 (48.6)	416.1 (49.7)	409.9 (49.3)
Overall	**397.0 (38.6)**	**406.1 (43.0)**	**410.2 (44.2)**	**406.5 (43.1)**

Timing Study

The time limit for the new SEE was set at four hours. The average total test time on the new SEE was about 164.1 minutes (2 hours, 44 minutes, 6 seconds), with a standard deviation of 43.8 minutes (43 minutes, 48 seconds). Total test times did not differ significantly based on students' Year in Program. Only 20 examinees ran out of time before completing the test. The rate of examinees exhausting their time prior to test completion is comparable to the old SEE.

Reliability

Reliability is a statistical psychometric indicator that represents the precision of test scores. Reliability is often described conceptually as the extent to which the scores are free of systematic error. The lower the error, the more reliable the scores, and so the more useful they will be to stakeholders. In the specific context of the SEE, more reliable scores will help educators and students better identify the specific domains that are areas of strength and weakness.

The reliability of domain-level information was computed for the SEE as a function of the standard error of measurement (SEM) and standard deviation (SD). Specifically, the formula is:

$$\rho = 1 - \frac{SEM^2}{SD^2}$$

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

Table 3. Score Reliability, Old SEE

Domain	**Reliability Index**
Professional and Legal Aspects	0.49
Anatomy Physiology & Pathophysiology	0.65
Pharmacology	0.62
Basic Principles of Anesthesia	0.45
Advanced Principles of Anesthesia	0.50
Overall	0.83

Table 4. Score Reliability, New SEE

Domain	**Reliability Index**
Basic Sciences	0.77
Equipment, Instrumentation & Technology	0.76
Basic Principles of Anesthesia	0.79
Advanced Principles of Anesthesia	0.77
Overall	0.93

The range of reliability is from 0 to 1, where higher values indicate higher reliability. Table 3 contains the score reliability for overall and domain-level scores for both the old SEE, while Table 4 contains score reliability for overall and domain-level scores for both the New SEE. While direct comparisons are not feasible due to the different content rubrics between the old and the new SEE, there is a clear improvement in score reliability, some quite substantial, for the new SEE. The reliability of the overall score improved from 0.83 to 0.93. While on the old SEE, the sub-score reliability ranged from poor to moderate (0.49 to 0.65), the sub-score reliabilities for the new SEE were more consistent across domains and were all reported above 0.75. This improved reliability should assure educators that the scores arising from the new SEE are more accurate indicators of their students' knowledge. Thus, the scores should be more useful for identifying students' strengths and weakness and for planning future study.

Predictive Validity

One of the goals of the SEE reconfiguration was to improve the predictive validity of the SEE in relation to NCE performance. In other words, it was desirable that the new SEE would help educators understand how students may perform on future attempts of the NCE. Ideally, the educators would be able to identify students who were "at risk" for failing the NCE.

A correlation study was undertaken to evaluate the predictive power of the new SEE. To date, 871 examinees have taken the new SEE, and have also gone on to take their first challenge of the NCE. This sample provides a basis for comparing performance on the SEE to eventual performance on the NCE. Figure 1 displays an X-Y plot ("scatter plot"), in which each point (+) represents an examinee's new SEE score (horizontal axis) plotted against their eventual NCE score (vertical axis). A clear, positive correspondence exists between the new SEE scores and the NCE scores.

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

The Pearson correlation between the two sets of scores was r = 0.62. If we focus on examinees who are close to graduation (program year three and above) and who reported (in a survey question) that they studied at least 30 hours for the SEE, the SEE:NCE correlation improves to r = 0.7. This result represents a moderately high degree of association between SEE performance and NCE performance. Nearly 50% of the variation in NCE scores can be explained solely by performance on the new SEE.

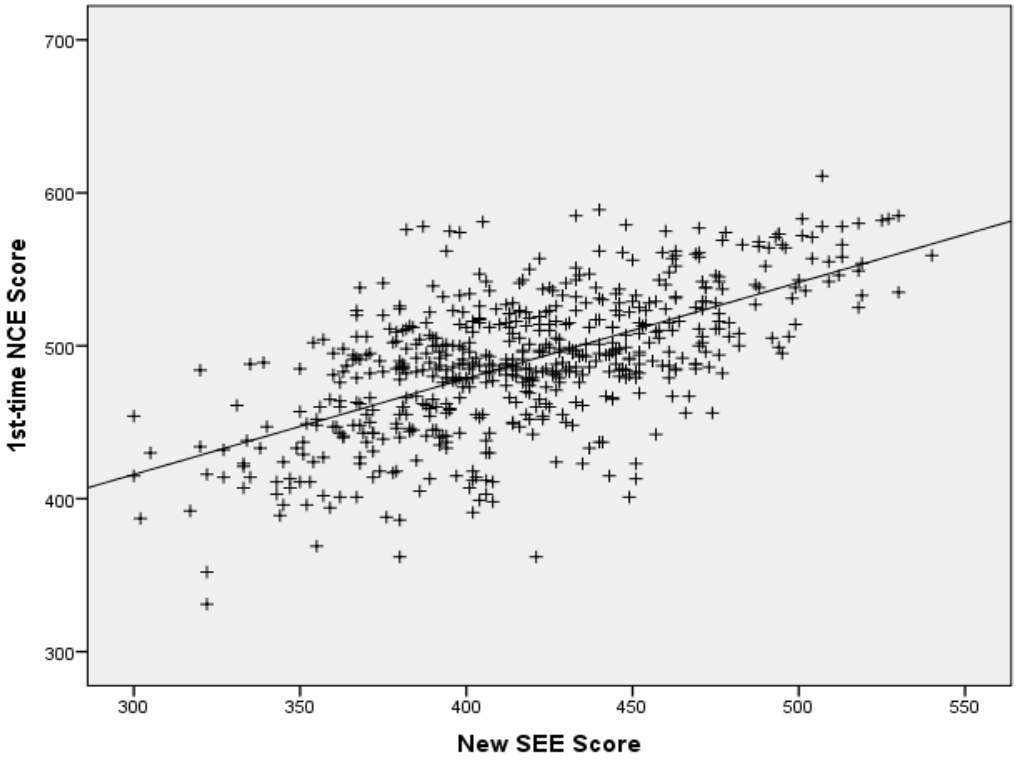

Figure 1. X-Y Plot of New SEE Scores vs. Eventual First-Time NCE Scores

Figure 2 displays two "box and whisker" plots, which can be understood as a "top-down" view of a normal (bell curve) distribution. The box plot on the left represent SEE scores for examinees who eventually failed their first attempt of the NCE, the plot on the right for those who went on to pass. The bold line at the center of each box represents the median (50th percentile) of the SEE score for either group. The top and bottom edges of the box represent the interquartile range (IQR), with the lower box-edge representing the 25th percentile, and the top box-edge representing the 75th percentile. The box thus represents the middle 50% of the scores in the distribution. The lines at the top and bottom of the "whiskers" represent 1.5 multiplied by the IQR, above and below the median. Clearly, the mean SEE score for the eventual NCE passers is higher than the mean for the NCE non-passers. Also, the degree of overlap between the two score distributions is quite small. The center of the passing group is over a full standard deviation above that of the failing group.

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

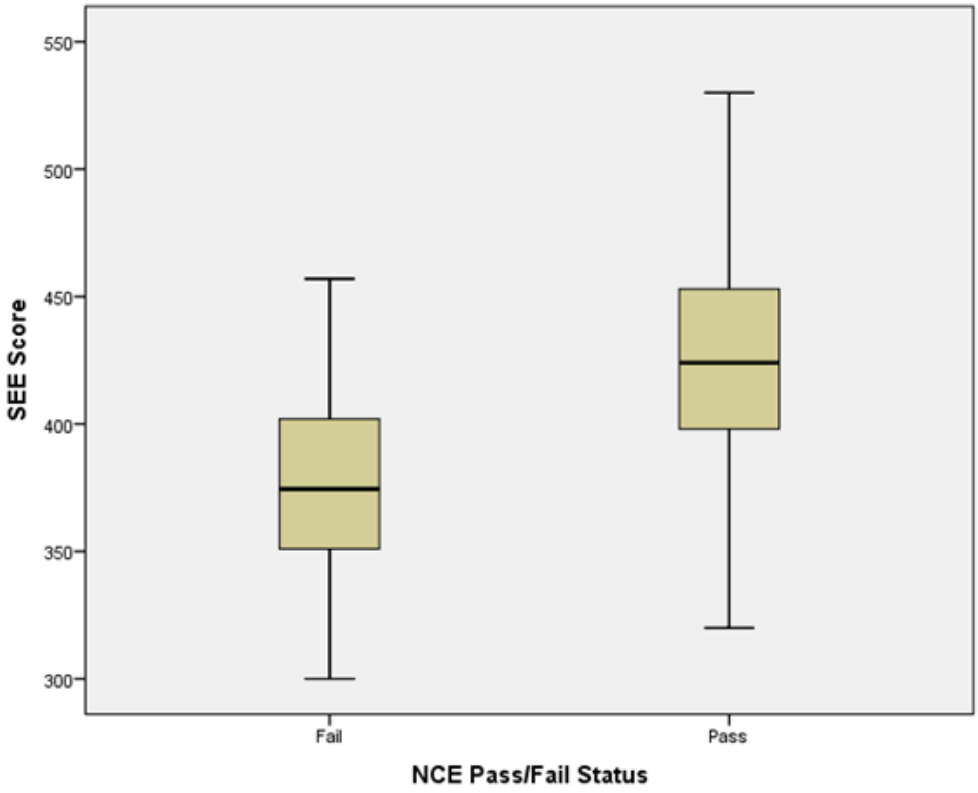

Figure 2. Box-and-whisker plots, SEE score distributions for those who went on to pass vs. fail first NCE attempt.

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

Figure 3 presents a sidelong view of the SEE scores for the NCE passing and failing groups. Again, there is a clear upward shift in SEE scores for students who eventually passed their first NCE attempt.

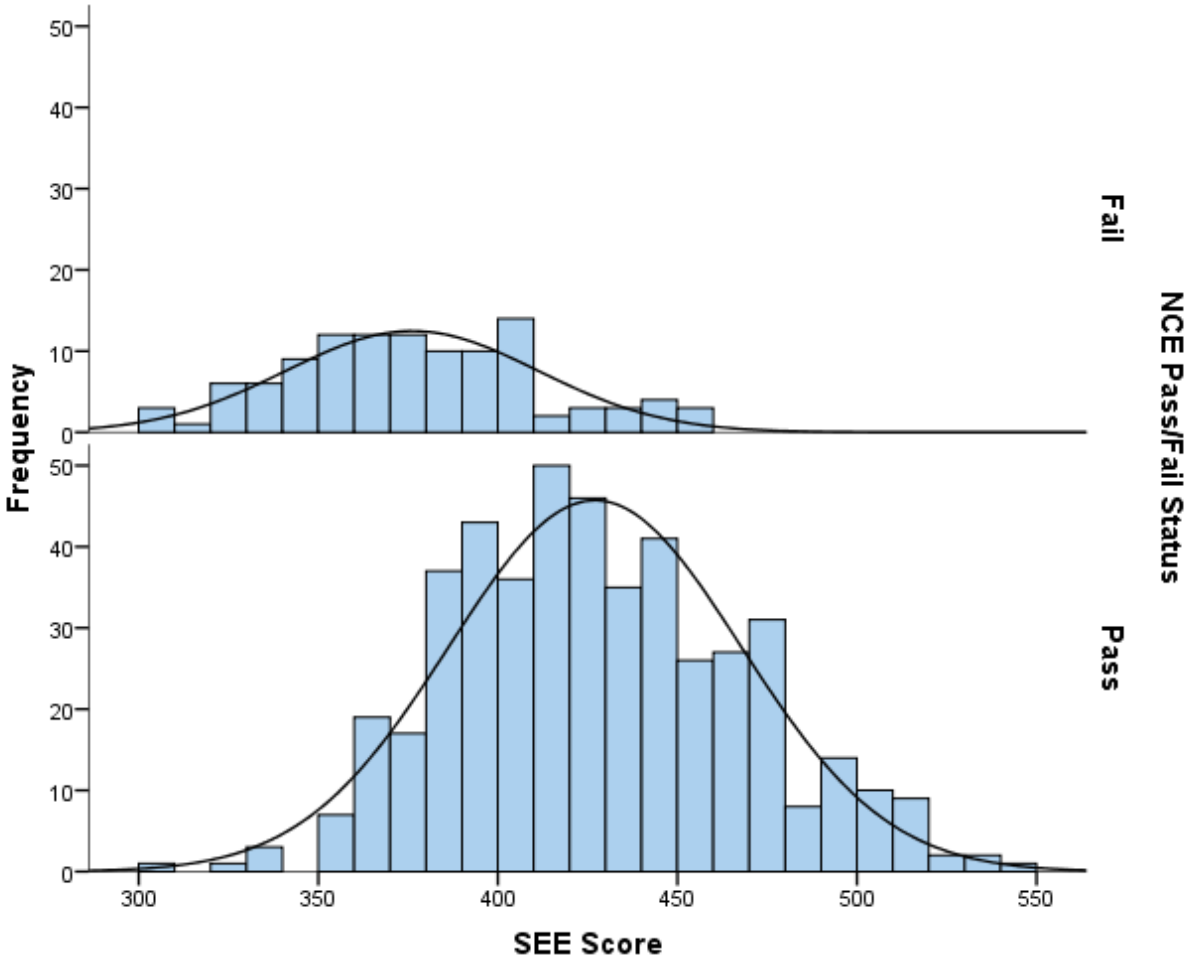

Figure 3. Normal distributions of new SEE scores for those who went on to pass vs. fail first NCE attempt.

Table 3 summarizes the new SEE score for both the NCE passing and failing groups.

Table 3. Summary of New SEE Scores by First-Time NCE Performance

	SEE Score	
NCE Pass / Fail Status	**Average**	**SD**
Fail	376.1	35.2
Pass	427.0	40.7

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631
NBCRNA
855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

Survey feedback

In addition to the analysis of SEE scores, it was desirable to study examinee perceptions of the new SEE. After completion of the examination, test takers were administered eight questions assessing their attitudes toward the revisions made to the SEE. The survey questions are presented below, along with the percentage of examinees who indicated agreement or disagreement with the statement. The following results represent the responses of 987 examinees, collected from 9/1/16 through 1/12/17. Overall, the results were positive and supportive of the new SEE Examination.

1. *In my exam, the areas of the SEE content outline were fairly represented.*

Answer Choices	Responses	N
Strongly Agree	**21.4%**	211
Agree	**66.8%**	659
Disagree	**10.0%**	99
Strongly Disagree	**1.1%**	11
No Response	0.7%	7
Total		**987**

2. *The questions on my test today fairly reflected the knowledge of the subject matter I have been taught in my educational program.*

Answer Choices	Responses	N
Strongly Agree	**16.3%**	161
Agree	**63.9%**	631
Disagree	**17.1%**	169
Strongly Disagree	**1.9%**	19
No Response	0.7%	7
Total		**987**

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

3. *The questions on my test today were clearly written.*

Answer Choices	Responses	N
Strongly Agree	**19.0%**	188
Agree	**64.2%**	634
Disagree	**15.1%**	149
Strongly Disagree	**0.8%**	8
No Response	0.8%	8
Total		**987**

4. *It is helpful to have the SEE based on the content outline for the National Certification Examination (NCE).*

Answer Choices	Responses	N
Strongly Agree	**46.6%**	460
Agree	**47.8%**	472
Disagree	**4.0%**	39
Strongly Disagree	**0.6%**	6
No Response	1.0%	10
Total		**987**

5. *Receiving SEE scores for the same domains that also appear on the NCE (e.g., Basic Sciences, Equipment, etc.) will help me identify relative strengths and weakness when I prepare for the NCE.*

Answer Choices	Responses	N
Strongly Agree	**56.6%**	559
Agree	**40.6%**	401
Disagree	**1.4%**	14
Strongly Disagree	**0.7%**	7
No Response	0.6%	6
Total		**987**

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

6. *Taking the SEE will help me to better prepare for the NCE.*

Answer Choices	Responses	N
Strongly Agree	**50.9%**	502
Agree	**42.7%**	421
Disagree	**4.3%**	42
Strongly Disagree	**1.4%**	14
No Response	0.8%	8
Total		**987**

7. *Taking the SEE gives me a sense of what to expect when I take the NCE.*

Answer Choices	Responses	N
Strongly Agree	**43.3%**	427
Agree	**50.8%**	501
Disagree	**4.2%**	41
Strongly Disagree	**0.8%**	8
No Response	1.0%	10
Total		**987**

8. *I had adequate time to complete my examination today (i.e., I did not feel rushed).*

Answer Choices	Responses	N
Strongly Agree	**43.5%**	429
Agree	**47.1%**	465
Disagree	**7.1%**	70
Strongly Disagree	**1.5%**	15
No Response	0.8%	8
Total		**987**

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax

Conclusions

Thus far, the new SEE has demonstrated it is meeting its intended goals, as evidenced by:

- Testing volume is about 28% higher on the new exam compared to a year ago.
- Overall scores on the new SEE are comparable to the legacy exam.
- The increased test time of four hours appears to be adequate for the clear majority of test takers to complete the examination.
- There is a moderately high, positive correlation between new SEE scores and eventual performance on the first NCE attempt, indicating marked improvement of the predictive validity of the SEE, especially for 3rd year students who devote substantial time preparing for the SEE.
- The reliability of scores on the new SEE, both overall and across all domains, have shown marked improvement, making domain-level scores a much sounder foundation on which to base remediation.
- Candidate perceptions of the new SEE are generally positive.

The NBCRNA will continue to collect data and to publish evaluation results along the dimensions of performance, testing time, predictive validity, examination reliability, and stakeholder perceptions for the new SEE.

# # #

8725 W. Higgins Rd.
Suite 525
Chicago, IL 60631

NBCRNA

855-285-4658
www.NBCRNA.com
708-669-7636 :: fax