

# **SEE Evaluation Report**

September 1, 2018 - August 31, 2019

## Table of Contents

<b>Introduction.....</b>	<b>1</b>
<b>Score and Timing Summaries, FY 2019.....</b>	<b>1</b>
<b>Reliability.....</b>	<b>2</b>
<b>Predictive Validity.....</b>	<b>3</b>
<b>Survey Feedback, FY 2019.....</b>	<b>7</b>
<b>Conclusions.....</b>	<b>11</b>

## Introduction

The fiscal year 2019 (FY 2019; September 1, 2018, through August 31, 2019) marks the third year since the NBCRNA launched a reconfigured version of the Self-Evaluation Examination (SEE) on September 1, 2016. Compared with the previous version of the SEE exam, the revised SEE featured the following changes:

- Alignment of SEE content outline with the four primary domains of NCE outline
- 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 to accommodate the increased test length
- Enhanced predictive capability with respect to future performance on the National Certification Examination (NCE)

The content outline for the SEE underwent an updating in May 2018 to reflect the revisions to the NCE content outline earlier that year.

The purpose of this report is to utilize accumulated SEE data from FY 2019 for the ongoing evaluation of SEE program.

## Score and Timing Summaries

From September 1, 2018, through August 31, 2019 (i.e., FY 2019), the SEE has been administered to 4,441 examinees. This testing volume represents a 6% increase compared to 4,190 in fiscal year 2018 (FY 2018) and 16% increase compared to 3,845 examinees in fiscal year 2017 (FY 2017). Table 1 contains the average total scores for the SEE in FY 2019, FY 2018 and FY 2017. The average scores are comparable<sup>1</sup>. The SEE average scores for all reported categories in this table are nominally higher in FY 2019.

### Score Performance

Table 1. Comparison of Average Scores, FY2017, FY2018 and FY2019 SEE by Year in Program

Year in Program	SEE Average (SD) FY2019 <sup>1</sup>	SEE Average (SD) FY2018 <sup>1</sup>	SEE Average (SD) FY2017
Year 1	404.8 (45.0)	393.0 (47.5)	393.0 (40.9)
Year 2	409.7 (43.5)	406.1 (44.1)	402.9 (42.9)
Year 3 and Up	424.7 (41.3)	419.7 (41.4)	410.3 (43.6)
<b>Total</b>	<b>417.6 (43.1)</b>	<b>411.1 (44.0)</b>	<b>405.3 (43.3)</b>

<sup>1</sup> While the SEE content outline was revised in May 2018, it is still possible to compare *overall* scores before and after the 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.

Tables 2 contain descriptive summaries for the overall score and domain-level scores by year in program. More detailed information, including percentile transformations, is available in the FY2019 SEE Interpretive Guides on the NBCRNA website.

**Table 2. SEE Overall and Domain-Level Scores by Year in Program, FY 2019 (September 1, 2018–August 31, 2019)**

	1st Year in Program		2nd Year in Program		3rd Year in Program		All	
	Avg	SD	Avg	SD	Avg	SD	Avg	SD
<b>Total</b>	<b>404.8</b>	<b>45.0</b>	<b>409.7</b>	<b>43.5</b>	<b>424.7</b>	<b>41.3</b>	<b>417.6</b>	<b>43.1</b>
Basic Science	407.5	51.9	407.2	50.8	418.1	48.8	413.1	50.1
Equipment, Instrumentation and Technology	405.8	49.5	413.4	48.7	429.5	45.3	421.8	47.7
General Principles of Anesthesia	404.9	50.2	410.0	47.0	424.2	46.2	417.4	47.3
Anesthesia for Surgical Procedures and Special Populations	404.9	47.1	411.5	49.0	430.3	47.5	421.4	49.1

### Timing Study

The time limit for the SEE was set at four hours (240 minutes). The average total test time on the SEE in FY2019 was 176.5 minutes (about 2 hours and 57 minutes), with a standard deviation of 44.0. A total of 31 examinees reached the time limit before completing all questions on the examination. Total test times did not differ significantly based on students' Year in Program. This average testing time is comparable to what was observed in FY 2018, 174.3 (44.0) minutes, and in FY 2017, 164.9 (44.0) minutes. The number of examinees running out of time is, however, fewer than those in FY 2018 and FY 2017, 41 and 36, respectively.

### Reliability

Reliability is a psychometric indicator which 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 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:

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

The range of reliability is from 0 to 1, with higher values indicating higher reliability. Table 3 contains the score reliability information for overall and domain-level scores for the SEE from September 2018 through August 2019.

Table 3. SEE Total and Domain Score Reliability FY 2019 (September 1, 2018 through August 31, 2019)

Domain	Reliability Index
Basic Sciences	0.78
Equipment, Instrumentation and Technology	0.76
General Principles of Anesthesia	0.76
Anesthesia for Surgical Procedures and Special Populations	0.78
<b>Overall</b>	<b>0.93</b>

Similar to those in FY 2017 and FY 2018, since the substantial revisions to the SEE in September 2016, the reliability of the overall scores improved from 0.83 to 0.93; the sub-score reliability increased as well. This improved reliability should help assure educators that the scores arising from the current SEE are accurate indicators of their students' knowledge. Thus, the assessment obtained by use of the SEE should be more useful for identifying students' strengths and weaknesses and for aiding in planning future study and/or remediation.

### Predictive Validity

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

A correlation study was undertaken to evaluate the predictive power of the SEE. Of the examinees who took the SEE in FY 2019, 2,650 have also gone on to complete their first attempt at the NCE, sometime between September 1, 2018 and December 10, 2019. This sample provides a basis for comparing performance on the SEE to eventual performance on the NCE. Figure 1 displays an X-Y plot (scatterplot), in which each point represents an examinee's SEE score (horizontal axis) plotted against their first-attempt NCE score (vertical axis). A clear, positive correspondence exists between the SEE scores and the NCE scores.

The Pearson correlation between the two sets of scores was  $r = 0.61$ . If we focus on examinees who had taken the SEE in their second or higher year of the program and at least three months prior to their first attempt at the NCE, the SEE–NCE correlation was higher at  $r = 0.64$ . This result represents a strong positive correlation between SEE performance and NCE performance. That is, about 41% of the variation in NCE scores can be explained solely by performance on the SEE.

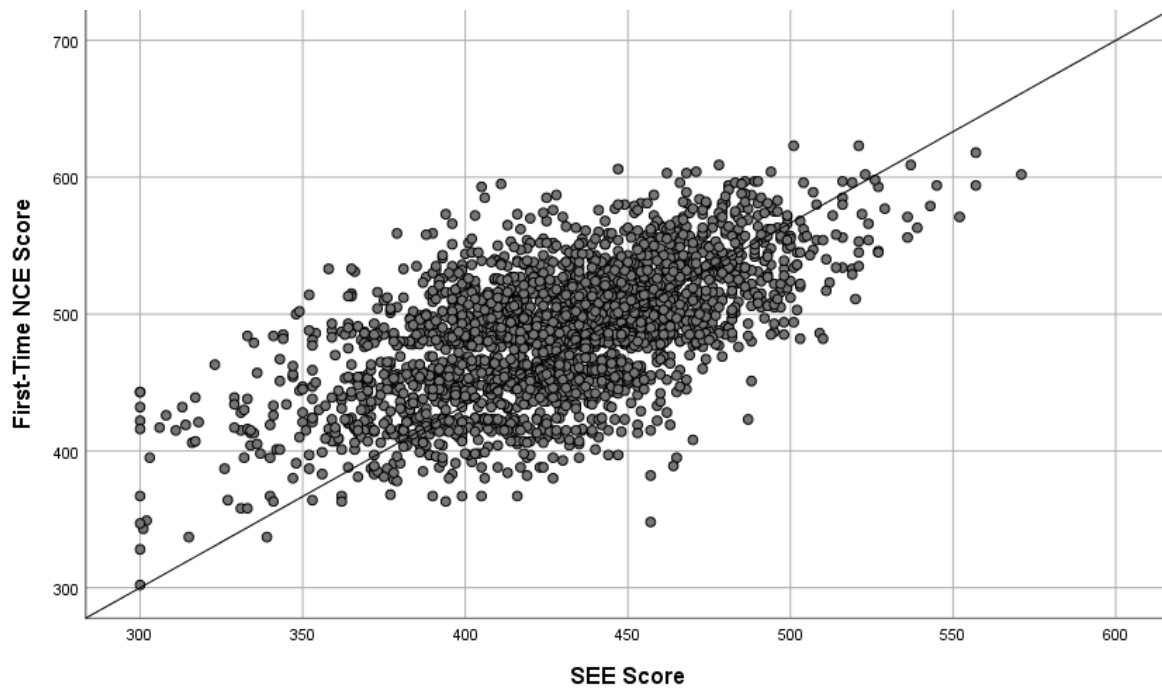


Figure 1. X-Y plot of FY 2019 SEE Scores vs. First-time NCE Scores

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 represents SEE scores for examinees who failed their first attempt of the NCE; the plot on the right represents those examinees who passed the NCE on their first attempt. The bold line at the center of each box represents the median (50<sup>th</sup> 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 25<sup>th</sup> percentile, and the top box-edge representing the 75<sup>th</sup> 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 times the IQR, above and below the median. The few dots below or above the lines signify extreme values. 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.

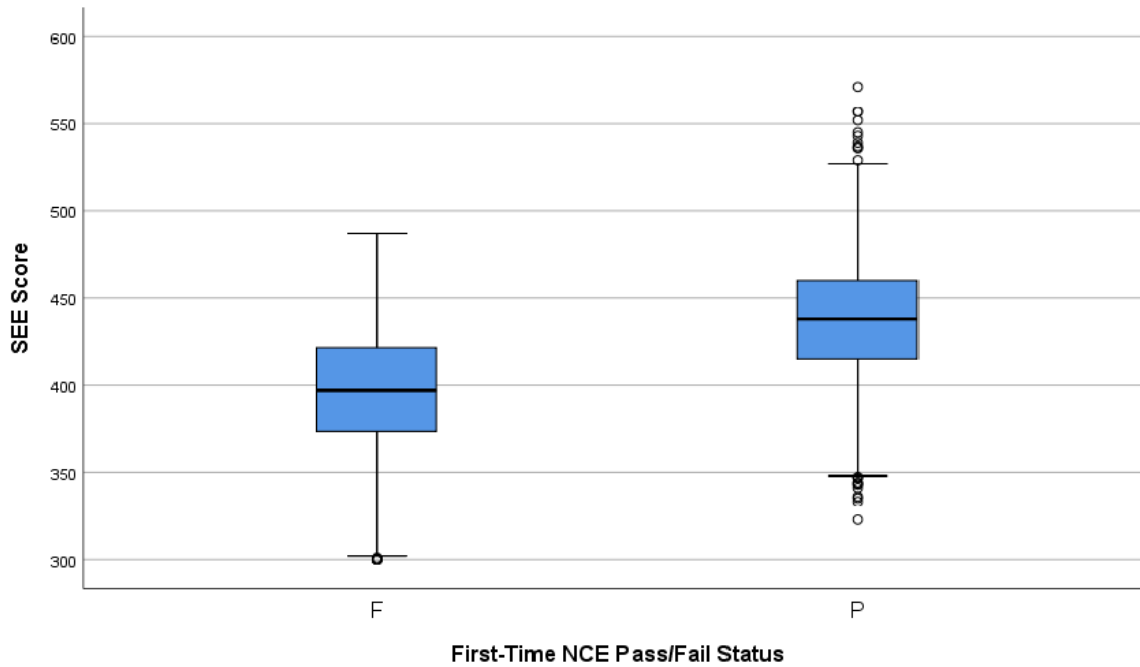


Figure 2. Plots of FY 2019 SEE Scores for Passing vs. Failing First NCE Attempt

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 NCE at their first attempt.

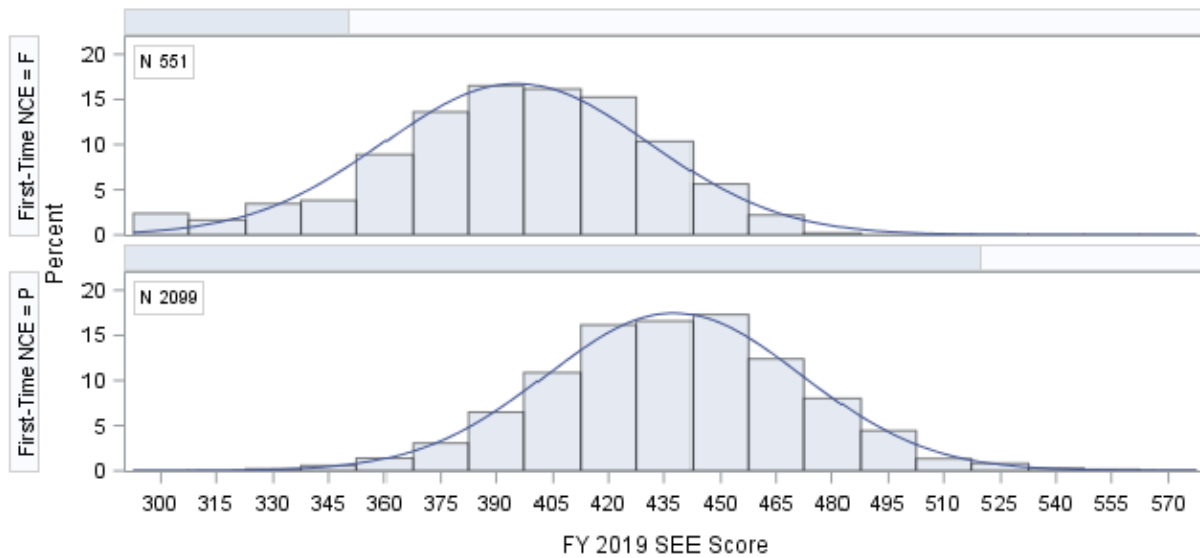


Figure 3. Normal Distributions of SEE Scores for Passing vs. Failing First NCE Attempt

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

Table 3. Summary of SEE Scores by First-Time NCE Performance FY 2019

<b>First-Time NCE Performance</b>	<b>N</b>	<b>SEE Score</b>	
		<b>Average</b>	<b>SD</b>
<b>Fail</b>	551	395.2	35.7
<b>Pass</b>	2,099	437.5	34.3



## Survey feedback

In addition to the analysis of SEE scores, NBCRNA sought examinee perceptions of the SEE. After completion of the examination, randomly half of SEE takers were surveyed about their attitudes towards SEE. The survey prompts appear below, along with the percentage of examinees who agreed or disagreed with them. The following results represent the responses of 2,167 examinees in FY 2019. For comparison, the responses of 2,074 and 1,970 SEE takers in FY 2018 and FY 2017, respectively, were also included for the same questions. Overall, the results in all three years are consistently positive and supportive of the SEE Examination.

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	<b>23.0%</b>	499	22.8%	473	22.1%	435
<b>Agree</b>	<b>64.0%</b>	1386	64.4%	1,336	64.9%	1,278
<b>Disagree</b>	<b>10.4%</b>	226	10.9%	227	11.3%	223
<b>Strongly Disagree</b>	<b>2.0%</b>	44	1.3%	27	1.0%	20
<i>No Response</i>	0.6%	12	0.5%	11	0.7%	14
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	<b>19.7%</b>	426	18.4%	382	16.5%	326
<b>Agree</b>	<b>62.2%</b>	1347	63.4%	1,314	62.9%	1,239
<b>Disagree</b>	<b>15.5%</b>	336	15.6%	323	18.0%	354
<b>Strongly Disagree</b>	<b>2.1%</b>	46	2.1%	44	1.9%	37
<i>No Response</i>	0.6%	12	0.5%	11	0.7%	14
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	<b>18.2%</b>	395	18.7%	388	18.3%	360
<b>Agree</b>	<b>65.2%</b>	1,413	65.6%	1,361	64.1%	1,263
<b>Disagree</b>	<b>14.9%</b>	322	13.8%	287	16.1%	317
<b>Strongly Disagree</b>	<b>1.2%</b>	27	1.4%	30	0.9%	17
<i>No Response</i>	0.5%	10	0.4%	8	0.6%	13
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	44.0%	953	<b>46.2%</b>	958	47.0%	926
<b>Agree</b>	51.5%	1,117	<b>49.0%</b>	1,017	47.8%	941
<b>Disagree</b>	3.0%	65	<b>3.2%</b>	66	3.7%	74
<b>Strongly Disagree</b>	0.9%	19	<b>1.3%</b>	26	0.4%	8
<i>No Response</i>	0.6%	13	0.3%	7	1.1%	21
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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 weaknesses when I prepare for the NCE.

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	<b>51.9%</b>	1,125	54.2%	1,125	55.3%	1,090
<b>Agree</b>	<b>44.1%</b>	956	41.5%	860	41.6%	820
<b>Disagree</b>	<b>3.0%</b>	64	2.8%	59	1.7%	35
<b>Strongly Disagree</b>	<b>0.6%</b>	12	1.0%	20	0.7%	13
<i>No Response</i>	0.5%	10	0.5%	10	0.7%	12
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	<b>44.9%</b>	973	47.5%	986	49.4%	973
<b>Agree</b>	<b>48.4%</b>	1,049	46.1%	956	44.3%	873
<b>Disagree</b>	<b>5.1%</b>	111	4.4%	92	4.4%	86
<b>Strongly Disagree</b>	<b>1.1%</b>	23	1.3%	26	1.3%	25
<i>No Response</i>	0.5%	11	0.7%	14	0.6%	13
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	<b>40.0%</b>	867	42.4%	880	42.8%	843
<b>Agree</b>	<b>52.5%</b>	1,137	51.0%	1,057	51.4%	1,013
<b>Disagree</b>	<b>5.9%</b>	127	4.8%	100	4.3%	85
<b>Strongly Disagree</b>	<b>1.0%</b>	22	1.2%	24	0.7%	13
<i>No Response</i>	0.6%	14	0.6%	13	0.8%	16
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

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

Answer Choices	FY 2019		FY 2018		FY 2017	
	Responses	N	Responses	N	Responses	N
<b>Strongly Agree</b>	42.7%	925	44.2%	917	44.1%	868
<b>Agree</b>	46.4%	1,005	45.6%	945	46.3%	912
<b>Disagree</b>	7.9%	172	7.7%	159	7.0%	138
<b>Strongly Disagree</b>	2.3%	49	2.1%	43	1.9%	38
<i>No Response</i>	0.7%	16	0.6%	13	0.7%	14
<b>Total</b>		<b>2,167</b>		<b>2,074</b>		<b>1,970</b>

## Conclusions

The analysis and evaluation of the FY 2019 SEE examination data showed results consistent with the FY 2018 and FY 2017 SEE examination data, confirming once again that the SEE has performed in a manner as intended based upon the stated goals. When the three-years of SEE examination data are compared, the main findings are:

- Testing volume continued increasing in FY 2019. From September 2018 to August 2019, the SEE has been administered to 4,441 examinees, representing 6% increase compared to 4,190 in FY 2018 and 16% increase compared to 3,845 examinees in FY 2017, which was 30% higher compared to the takers of the previous version of the SEE in FY 2016.
- Overall scores, as well as content domain-specific scores for the SEE from FY 2019, FY 2018 and 2017, are comparable, and slightly higher for FY 2019.
- The increased test time of four hours (240 minutes) appears to be adequate for the majority of test takers to complete the examination.
- There is a strong positive correlation between SEE scores and performance on the first NCE attempt, especially for second-year and third-year students who take the SEE at least three-months prior to their first NCE exam.
- The reliability of scores on the SEE in FY 2019, both overall and across all specific domains, were comparable to those in FY 2018 and FY 2017. In addition, it is important to point out that the reliability of SEE scores for 2017, 2018 and 2019 were substantially improved in comparison to the previous SEE format, making domain-level scores a much sounder foundation upon which to base remediation.
- Candidate perceptions of the SEE in FY 2019, FY 2018 and FY 2017 were generally positive and the satisfaction was even higher in FY 2019 and FY 2018 regarding the content relevance.

The NBCRNA will continue to collect examination performance data and publish evaluation results for the communities of interest. This data will include testing time, predictive validity, examination reliability, and stakeholder perceptions to guide the effective purposeful planning and utilization of the SEE examination.