THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

NBSTSA Comprehensive Guide to Professional Certification Exam Development

This guide is provided by the National Board of Surgical Technology and Surgical Assisting (NBSTSA), which administers the Certified Surgical Technologist and Certified Surgical First Assistant examinations. The NBSTSA is accredited by the National Commission for Certifying Agencies (NCCA), and our examinations comply with the NCCA Standards.

This guide will demonstrate the steps in the development of the outline that defines the contents of our examinations (CST® and CSFA®), the process of writing test items, item review, how items are pretested, examination content review, and ultimately, the launch of new examination forms each year.

Psychometrics is the science of measuring mental capacities and processes, and our certification examinations are aimed at measuring individual skills and knowledge at the entry level for the professions of surgical technology and surgical first assisting. We work with our psychometricians and test developers at PSI (a third party), and with our subject matter experts (SME) to develop valid test items that measure what's important. We use psychometrics to assess those items, assemble tests, and establish valid examination passing scores.

Significantly, our test development processes are based on scientific rigor and well-established standards in the professional examination industry, which ultimately helps to ensure public safety.

The entire process is overseen by the NBSTSA Board of Directors, NBSTSA's leadership, which consists of 10 individuals appointed by NBSTSA, the Association of Surgical Technologists, and the American College of Surgeons, including a surgeon and a member representing the interests of the public. The Board of Directors includes CST® and CSFA® practitioners and educators in both fields.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

Test Development Overview

A person new to the world of credentialing might think, "How hard could it be to create a valid credentialing examination for surgical technologists or surgical first assistants?" Some might think that it would just be a matter of getting a couple of people to write test items, put them on a test, and then we are ready to go, possibly forever. But creating a quality examination is much more complicated and requires a standardized process to create an examination that supports the purpose of our credentials and is also legally defensible.

NBSTSA's credentialing examinations are used to make high-stakes decisions. A high-stakes decision is one that has meaningful consequences for the candidate and/or the public.

For instance, test scores from our credentialing exams are often used to make decisions as to whether an individual:

- Is allowed to practice surgical technology or surgical first assisting in an operating room or surgical center.
- gets hired, promoted, or receives a pay raise, or
- is deemed competent and safe enough to practice in each state or with a given employer.

We recognize that whether individuals obtain a credential can significantly impact their lives and sometimes the welfare of surgical patients.

Therefore, it is important the test scores from credentialing exams are valid, reliable, and legally defensible.[1]

[1] "A Psychometrician's Guide to Smart Test Development." Kryterion. Kryterion, https://doi.org/https://www.kryterion.com/psychometricians-guide-smart-test-development/#.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

Reliability

It is important that NBSTSA's certification programs provide test scores that can be considered sufficiently reliable and valid to withstand a challenge. Our tests should be developed in such a way that if a candidate took the same exam two times in a row, without any kind of remedial action between test takes, that candidate would score similarly on both takes.

Validity

NBSTSA strives in our test development processes for validity in our examinations. That is, we strive to measure what we claim to measure. Test items must be closely linked to the focus of our examinations. NBSTSA examinations measure job-related competencies for surgical technologists or for surgical first assistants.

Our test development processes have the goal of ensuring content validity. That is, we use our process to make sure that our test items are directly connected to job-related tasks in surgical technology and surgical first assisting.

Our examinations' content validity measure is estimated by having multiple groups of subject matter experts (SMEs) in these fields review our test items. These SMEs are given the test items and the test content outline, and any items the SMEs identify as being inadequately matched or flawed in any way are either revised or dropped from our examinations.[2]

NBSTSA Subject Matter Experts – Committee Process

NBSTSA convenes committees of subject matter experts (SMEs) for the different steps in the process, including:

- Job Analysis
- Standard Setting
- Item Writing
- Item Review (ERC)
- Examination Review (ERC)

[2] "Test Validity." Criterion. Professional Testing, https://doi.org/https://www.proftesting.com/test_topics/test_quality_validity.php.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

These committees include practitioners and educators from the field, and each of the four committees above consists of two separate committees, one for each examination. The number of members of each committee varies for each group. Care is taken to include beginning practitioners, those with experience, and a wide range of backgrounds and geographic locations on each committee. We attempt to minimize crossover in membership from one committee to the next to maximize the perspectives and input provided.

Examination Development

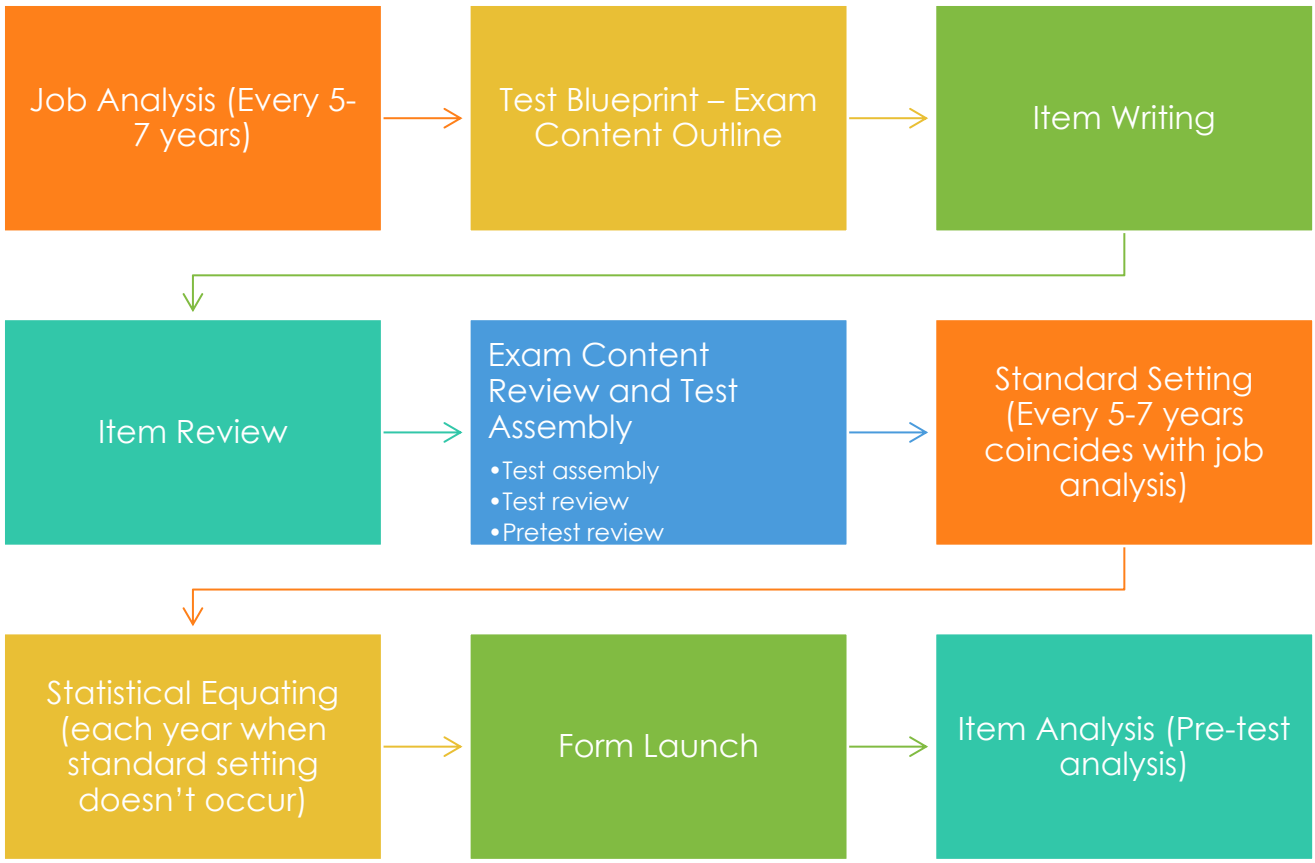
The examination development process, from item writing through the launch of new forms, takes place each year, and committees meet in person throughout the year, completing the process in December to launch test forms for the coming year.

The exam development and launch process is a 7-step process:

1. Job Analysis (Every 5-7 years)
2. Test Blueprint – Exam Content Outline
3. Item Writing (based on the Exam Content Outline)
4. Item Review
5. Exam Content Review and Test Assembly
    a. Test assembly
    b. Test review
    c. Pretest review
6. Standard Setting (Every 5-7 years coincides with job analysis)
7. Statistical Equating (each year when standard setting doesn't occur)
8. Form Launch
9. Item Analysis (Pre-test analysis)

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

NBSTSA Examination Development Life Cycle – Step By Step



1. Job Analysis

Job analysis is a systematic process of gathering information regarding the duties required of a job and the human characteristics necessary to successfully perform those duties. This process is the first step in the exam development process. It involves defining the exam and identifying domains, tasks, and knowledge and skill statements that will be used to create the draft examination blueprint.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

The job analysis translates practice into a usable format for test development. The job analysis outlines and delineates the tasks and knowledge necessary for competent practice. A well-conducted job analysis helps ensure that a certification exam is related and relevant to the job and, thus, that the exam has content validity.

NBSTSA convenes a Job Analysis Committee of SMEs to help our psychometricians and test developers design the survey and help with the analysis of the study. The result of this survey and all the job analysis work is compiled into a final Job Analysis Report, and the executive summary of this report is always made available for public review.

A representative group of professionals nationwide at all experience levels is surveyed to determine what they believe a newly certified professional should know to perform their job competently. NBSTSA surveys thousands of individuals nationwide each time they undertake a job analysis process to ensure adequate data is received for analysis.

This survey is conducted to validate the task statements developed by the committee. The survey results provide a sound and objective basis for a validated assessment that accurately reflects the competencies required for competent job performance.

NBSTSA performs a new job analysis once every 4-5 years. The job analysis process includes:

- Appointment of Job Analysis Committee of SMEs
- Job Analysis kick-off meeting with Survey design
- Carrying out the survey (several months)
- Survey results compilation and analysis
- Job Analysis Committee meeting to review results and approve Examination Content Outline

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

2. Test Blueprint – Exam Content Outline

The examination content outline, the blueprint, is developed based on the job analysis. This blueprint reflects the most current knowledge and skills required for entry-level jobs and to ensure patient safety. It is a high-level knowledge area that is essential to the practice of the profession. The blueprint guides the item development and examination assembly process and ensures that the examination reflects the relative importance of the required knowledge and skills[1].

The examination content outline shows the test candidate precisely what to expect. Because of the secrecy with which organizations like NBSTSA keep live test items, many myths exist about what is on the exam. The truth is that every exam form matches the examination content outline exactly, and each test-taker may expect to receive the same number of items in each area of the test as described in the examination content outline, which is always available on the website at www.nbstsa.org.

At the end of this process, we go through an entire test development process based on the new examination content outline, including item writing, item review, exam review, etc.

3. Item Writing

Item writing involves creating and validating new items. This activity is conducted with a panel of SMEs for each of our examinations and each meeting lasts several days. All SMEs are specially trained in quality item writing and how to avoid common pitfalls when writing items. Our item writers are provided the current references from our accepted reference list so items can be properly referenced as they are written.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

All test items should

- Be categorized with a section of the Exam Content Outline.
- Assess something that is relevant and important to the job.
- Describe realistic and practical situations.
- Not use language verbatim from a reference.
- Include the most current information available.
- Avoid jargon and overly conversational language.
- Objectively describe individuals.
- Avoid terms that may perpetuate stereotypes.
- Avoid pronouns and instead use a title or description.
- Include words and phrases that are universally understood.
- Include technical terms but avoid regional or cultural terms, for example, instrument names that are regional only.

Identifying the cognitive level of an item will require consideration of the type of thinking required of a typical examinee responding to the item. NBSTSA currently uses a two-level classification system to identify the level of thinking needed to respond to an item. The models used are based on Bloom's taxonomy. The two-level system includes Recall and Application. The primary goal in classifying items to cognitive levels is to minimize or control the number of items classified to the first level such that the exam content is not overly rudimentary.

Level 1 – Recall

Recall items primarily test the recognition or recall of information. Such items require predominantly an effort of memory. They include the recall of specific facts, concepts, principles, processes, procedures, or theories. To simplify, such an item will ordinarily be asking: "What is X?"

Example: What is the most populous city in the United States?

A. Los Angeles
B. Chicago
C. Houston
D. New York

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

Level 2 – Application

Application items primarily test the interpretation or evaluation of information. These items require applying known or presented information to solve or address situations, problems, and scenarios. Items could require examinees to make judgments concerning the effectiveness, appropriateness, or best course of action for a particular situation. Items at this level will ordinarily be asking: "Given the following information about Y, what is X?"

Example: Which diagnosis is most appropriate for a 28-year-old individual who exhibits anhedonia, difficulty making decisions, and a recent change in sleeping habits?

A. Generalized Anxiety Disorder
B. Major Depressive Disorder
C. Dissociative Identity Disorder
D. Brief Psychotic Disorder

SMEs may encounter many pitfalls during item writing, and we utilize the services of testing professionals and psychometricians to help review items as they are written.

All items are then reviewed and edited by other SMEs[1]. Each item should measure only a single objective[1]. Each objective, however, should be measured by one or several items, depending on the test specifications[1].

After questions have been written and initially reviewed, they are moved to the next step, which is item review by a different committee of SMEs, our Examination Review Committee for the CST (CST-ERC) or the Examination Review Committee for the CSFA (CSFA-ERC), respectively.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

4. Item Review

After the items have been written, they undergo a thorough review process to ensure their validity and relevance. This review is done by SMEs. The items are reviewed and edited to ensure they conform to the test specification, are clear and unambiguous, and meet high-quality standards.

5. Exam Content Review and Test Assembly

Exam Content Review

The exam content is reviewed to ensure it accurately reflects the knowledge and skills required for the profession. This review is led by SMEs. The review process ensures that item content is accurate and reflects current practice. All approved items are reviewed periodically for accuracy, currency, and relevance.

Test Assembly

Test assembly involves putting the items our SMEs have written, reviewed, and approved for use on test forms. But it's not as simple as that.

Any test form must:

- Accurately reflect NBSTSA's test specifications (be valid)
- Be equivalent to any alternate test forms (be reliable)
- Not unfairly disadvantage any test takers (be fair)

Certification programs use one of two options for test form assembly, depending on the profession they represent, the number of test takers, and specific program needs. Currently, NBSTSA uses both options:

Fixed forms – where all test takers see one of a small number of alternate and equivalent test forms. Items and distractors are scrambled, and for security, care is taken that test takers in a group receive different forms.

Linear On the Fly Testing (LOFT) – where every test taker sees a unique and equivalent test form with different items. These items are selected from a pre-

approved pool of items that meet established content and psychometric parameters.

The LOFT option is one that NBSTSA has recently launched to enhance examination security. Using this option, we can use our one large pool of live test items to create, for example, 100 examination forms from a pool of 500 items. **Note**: these numbers are only presented as an example and don't represent NBSTSA's actual item pool size or the number of items we have and use. Those numbers are kept confidential for examination security purposes.

Test review

Test review by a committee of SMEs is an important validation step for our item pools and test forms.

When we use a fixed form, our ERC committee will thoroughly review each test form to ensure they are valid and fair.

When we use LOFT, the committee reviews all the items in the pool, not each form. This can save the committee and its SMEs time because they don't need to review the same item multiple times if it appears on multiple forms. Any items with content issues are removed from the LOFT pool at the test review meeting, reducing the need for an additional review after the meeting.

Pre-Test Review

An important part of the exam content review process that takes place during the meetings of the examination review committees is a review of items that have been going through the pre-test process. These items have been answered in a real-world setting by real test-takers as items that don't count for or against the test-taker.

Psychometricians present performance statistics on these items to the subject matter experts at this point, make recommendations to remove poorly performing items, or the SMEs may edit the items and put them back into the review process to potentially have a valid and fair item to use on the live examinations later.

This is an important step to ensuring that only items that are properly formed, are not trick items, are relevant to the content outline, and are fair end up being used as a "live" item on our examinations.

The reader of this guide will find more information about this process later in this guide.

6. Standard Setting

After the job analysis and one round of test development based off the updated examination content outline has been completed (Steps 4 and 5 below), another separate committee of SMEs is convened, the NBSTSA Standard Setting Committee.

Standard setting, or cut score setting, is a critical step in the test development process, which occurs after job analysis. It involves determining the minimum score a candidate must achieve to be considered competent in the skill or knowledge area the test is measuring. In simple terms, this is setting the passing score for the test.

The NBSTSA standard-setting meetings usually take one day per examination and occur in person. Standard-setting meetings occur only at the end of our job analysis cycle process, only once every 4-5 years.

The process begins with a panel of subject matter experts (SMEs) who are thoroughly familiar with the test content and the target population's abilities. For NBSTSA, these experts are always asked to remember that our examinations are for entry-level professionals.

These SMEs review each test item and judge the item's difficulty level. They estimate the probability that a minimally competent candidate would answer the item correctly. This process is always facilitated by a psychometrician who ensures that the panel's judgments are statistically sound and defensible.

Once the SMEs have rated all the items, their ratings are aggregated to form an initial recommended cut score based on an analysis of the statistical data provided by the ratings. This score is then reviewed and adjusted as necessary based on a variety of considerations, such as the consequences of false

positives (passing candidates who are not truly competent) and false negatives (failing candidates who are competent), the need for consistency with other tests in the same program, and legal and regulatory requirements.

The NBSTSA Board of Directors meets with the psychometrician, and any other necessary test development team members, to be presented with the findings of the standard-setting committee after the standard-setting meeting has been completed. The Board has final approval and oversight of the passing score (cut score) for all NBSTSA examinations.

Once the cut score is set, it is announced concurrently with the launch of the new examination forms.

After the cut score is set, it's not the end of the process. The cut score should be periodically reviewed and possibly adjusted to account for changes in the test content or the abilities of the test-taking population. This ensures the examination remains both valid and fair over time. NBSTSA typically only changes the cut score during standard setting following an examination development cycle after a job analysis.

Standard setting is a complex, iterative process that requires a combination of expert judgment and statistical analysis. Ensuring that certification tests are fair, valid, and valuable in their intended contexts is crucial.

NBSTSA convenes a Job Analysis Committee of SMEs to help our psychometricians and test developers design the survey and help with the analysis of the study. The result of this survey and all the job analysis work is compiled into a final Job Analysis Report, and the executive summary of this report is always made available for public review.

7. Year to Year - Equating

Equating takes place each year that standard setting doesn't occur

The reader of this guide might ask since we do a standard setting to set the passing score on the exam only once every 4-5 years, yet we develop new exam forms each year without undergoing the full review of a standard setting;

how do we keep the examination at the same level and fair for test takers from one year to the next?

We use a process called equating each year to ensure the difficulty level of the exams and all the different forms of the exams stay consistent. This maintains the fairness and validity of the exam over time when new forms of the test are introduced or when the abilities of the test-taking population change.

Equating is a statistical process used to ensure the difficulty level of different forms of an examination is consistent. This process is crucial in maintaining the fairness and validity of the assessment over time, especially when new forms of the test are introduced or when the abilities of the test-taking population change.

The Modified Angoff method is often used in the equating process. This method involves a panel of subject matter experts (SMEs) familiar with the test content and the abilities of the minimally competent candidate. These experts review each item on the new form of the test and estimate the probability that a minimally competent candidate would answer the item correctly. These ratings are then compared to the ratings from the original form of the test.

The equating process involves several steps. First, the SMEs' ratings for the new form are aggregated to form an initial cut score. This score is then adjusted based on the cut score of the original form of the test. The goal is to ensure that a candidate who scores at the cut score on the new form would have also scored at the cut score on the original form, even though the items may differ.

After the equating process, the new test form can be administered to candidates. The scores they receive will be comparable to those from the original form of the test, ensuring fairness and consistency.

The examination form equating process, particularly when using the Modified Angoff method, is a complex but necessary step in maintaining the validity and fairness of an examination over time. It ensures the scores are consistent and meaningful regardless of when or which form of the test a candidate takes. This process is typically conducted in subsequent years after an examination cut score is set and before a new job analysis occurs. It's an integral part of the lifecycle of a standardized test.

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

8. Launching New Forms

NBSTSA Launches New Forms in January, Each Year

Once the exam content has been finalized and reviewed, new exam forms are launched. This involves administering the exam to candidates and managing the scoring and reporting of results[1].

NBSTSA launches new test forms each January.

9. Pre-test Analysis

Pretesting is an essential step in the exam development process. During pre-test analysis, brand new items are included in a test as unscored items. This provides an opportunity to analyze item data before the item is added to the scored item pool. It involves testing the items on a representative sample of the population for whom the exam is intended. This helps to identify any potential issues with the items before they are included in the final exam. Pretesting items allows for gathering statistics regarding candidate performance on each new item. Evaluating pretest statistics allows for confirmation that newly developed items are performing within acceptable statistical parameters before the item can affect a candidate's exam score.

There are 175 items on each of the NBSTSA examinations, both the CST, and the CSFA. Of the 175 items, 150 are live items that are scored, and 25 are pre-test analysis items that don't count for or against the test-taker. These items have been written, reviewed by a separate committee of SMEs, and initially analyzed by our test development professionals, but we are gathering data on the performance of the item before it is added to the scored item pool. This allows us to ascertain item validity, how the item performs with real-world test takers, and to analyze item data, which helps to ensure items with problems are weeded out before the item is ever presented to an individual in a setting where it would count for or against the test taker.

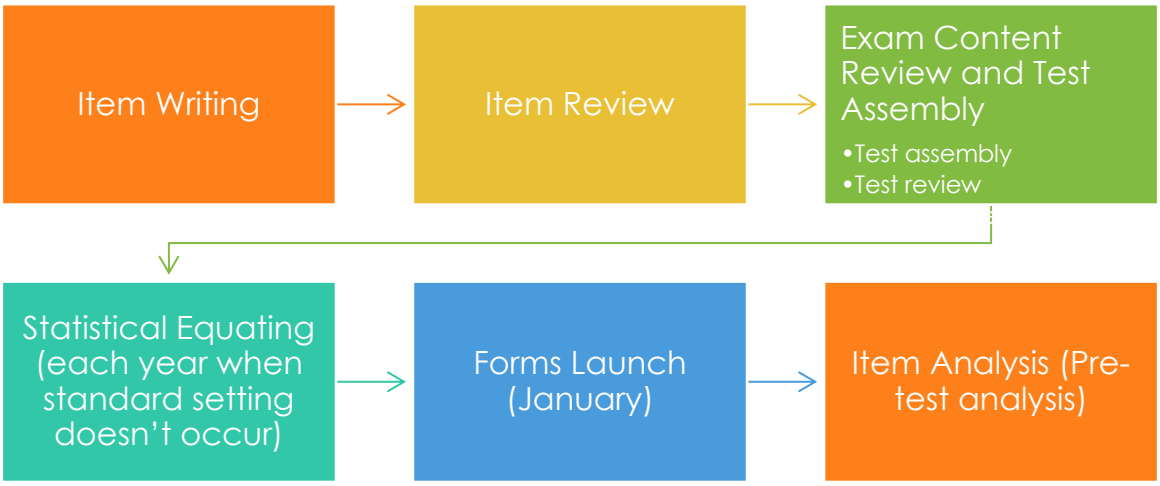
Steps Taken Each Year

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

In the year a job analysis has just been completed, all the steps above are performed, except equating is unnecessary because a standard setting takes place.

In the interim years, all the steps from 3 through 5, and 7 are completed by NBSTSA, including item writing, item review, pretest analysis, exam review, and test assembly, equating, and launching new forms all take place.

NBSTSA Annual Examination Development Steps
(Non Job Analysis/Standard Setting Years)



THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

Summary

NBSTSA launches new test forms each year after the above process that begins with item writing and ends with test review has been completed. The entire process, from item writing to test review, usually starts in January and ends in December, with new forms launching the next month.

Every 4 to 5 years, we undertake a job analysis and standard setting, a more comprehensive process beginning with a large-scale survey of certified and non-certified practitioners nationwide.

In the years between, we continue to develop new test items and launch new forms each year, utilizing an equating procedure to maintain the fairness and consistency of our exams from year to year.

This guide explains the steps involved in NBSTSA's exam development process: job analysis, item writing, item review, pre-test analysis, exam content review, test assembly, and launching new forms. Our guide also describes the psychometric methods and standards used to ensure our examinations' validity, reliability, and fairness.

This guide demonstrates how the NBSTSA examinations reflect the current knowledge and skills required for entry-level jobs in surgical technology and surgical first assisting.

The NBSTSA ensures that all these steps are followed in developing their certification exams, thus promoting quality patient care in the surgical setting[1]. The certification process is voluntary, although it is increasingly required by state law as well as most hospitals in the United States require certification as a condition of employment.

Holding and maintaining the CST or CSFA certification (or both) demonstrates the individual's mastery of a broad range of skills related to surgical procedures, aseptic technique, and patient care[1]. It is a means for upward mobility, a condition for employment, a route to higher pay, and a source of recognition nationwide.

Additional Resource(s)

THE NATIONAL BOARD
OF SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

1. NBSTSA
2. PSI Test Development Guide
3. NCCA 2021 Standards for the Accreditation of Certification Programs
4. Innovative, secure test delivery you can trust - PSI Exams
5. Test Development and Design With PSI
6. Test Development, Multi-Modal Delivery & Security | PSI
7. SUGI 25 – Test Development
8. A Psychometrician's Guide To Smart Test Development
9. American Psychological Association Standards for Educational and Psychological Testing