

**YARDSTICK WHITEPAPERS:**

**THE ASSESSMENT  
LIFE CYCLE**

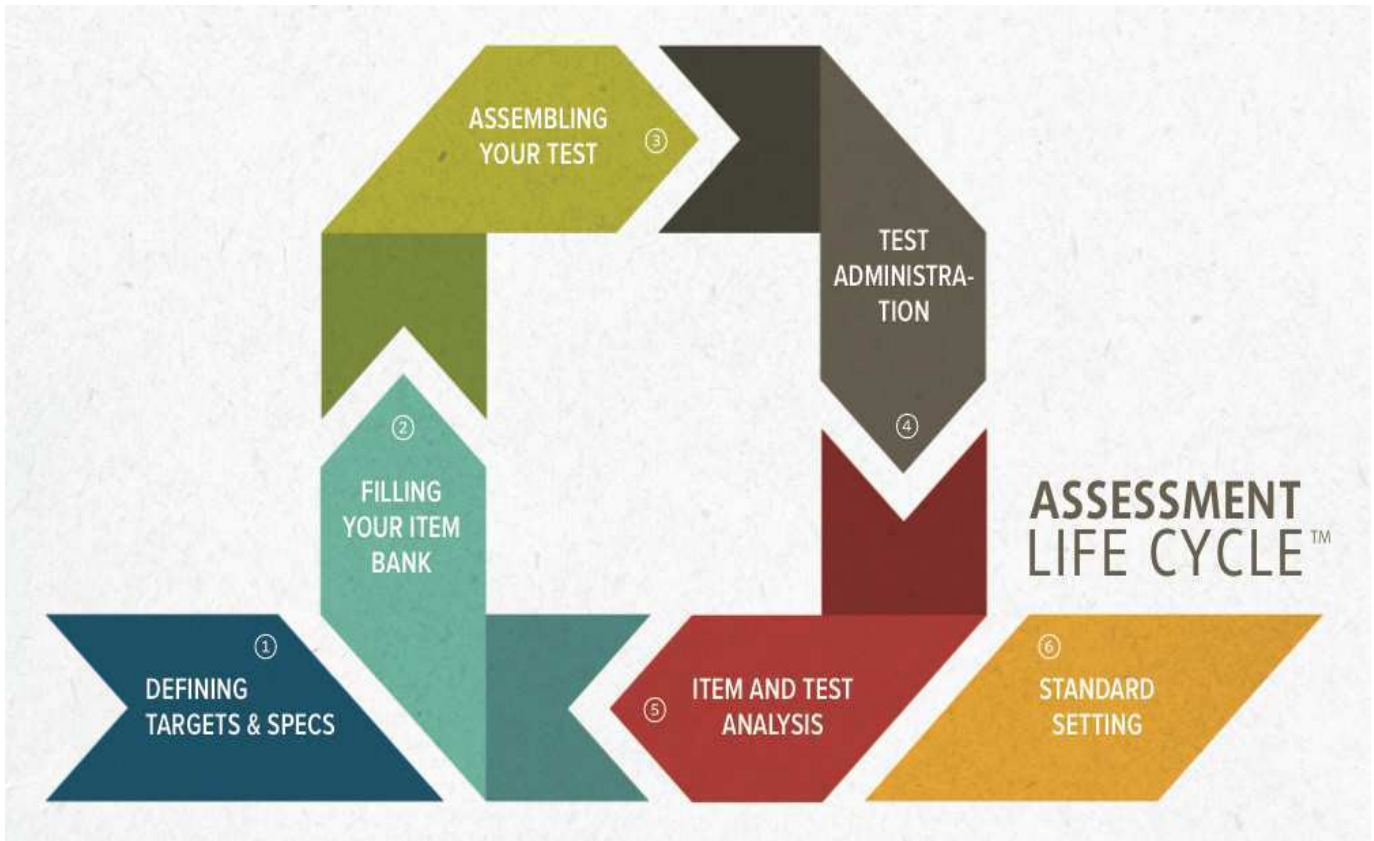


## TABLE OF CONTENTS

Introduction .....	3
Definition of roles .....	5
Stage 1: Defining purpose, targets, and creating specifications .....	9
Stage 2: Filling your item bank .....	10
Stage 3: Assembling your test .....	13
Stage 4: Test administration .....	14
Stage 5: Item and test analysis .....	16
Stage 6: Standard setting .....	18
Closing .....	19
List of psychometric services offered in each Assessment Life Cycle Stage .....	20

## INTRODUCTION

The Assessment Life Cycle is a way of organizing the processes involved in creating sound assessments into a series of easy-to-understand stages. The life cycle looks like this:



The first stage in the cycle, Defining Target and Creating Specifications, is where any assessment program starts. Here test developers decide what they'd like their assessment to be able to conclude about test takers, and finish by creating a crystal clear blueprint on the number and types of test questions that need to be written.

Stage Two, Filling Your Item Bank is about how to create test questions to fill out all the cells of your blueprint. What's most important here is that questions are constructed so they actually require the skill or knowledge indicated in the blueprint.

How to construct an examination is the essential part of Stage Three. Sometimes this is as easy as finding the right items from the right subject areas, but often, and particularly for certification exams, this step requires meticulous effort!

Test Administration conditions are the subject of Stage Four. The central concern here is to create conditions that are consistent with the 'seriousness' of the exam.

Stage Five is Item and Test Analysis. In this stage, actual data on how test takers responded to each item are analyzed to determine whether the item is testing what was intended, and is contributing to the quality of the test.

Lastly, in Stage Six, Standard Setting describes the method by which a passing score on the exam is set, and how well it represents the purpose, defined way back in Stage One.

When an examination program is well designed, each of the Stages represents a link in a chain. The overall strength of the chain determines how strong the claim is that passing the test is equivalent to possessing a sufficient amount of the skills and knowledge required. As always, any chain is only as strong as its weakest link.

Although following the Assessment Life Cycle is a necessity for high stakes testing programs (i.e., certification and licensure) elements of the Assessment Life Cycle can be applied to all testing contexts, regardless of stakes, in order to help make testing defensible and fair for candidates.

This white paper will introduce you to the Assessment Life Cycle and provide the foundation of knowledge to begin planning how to ensure the defensibility of your testing program. We will elaborate on each of the stages providing a solid framework of understanding for all of the processes and procedures.

Future white papers for each stage of the Assessment Life Cycle will take the level of detail beyond what can be covered in this introductory document, building on the foundation of knowledge.

## DEFINITION OF ROLES

Before we begin delving into the details of each stage of the Assessment Life Cycle it is important to define the roles of each of the players in the test development enterprise. Building and maintaining an examination program is a time and resource intensive exercise requiring people with many different types of expertise. Depending on the size of your assessment program one person may take on many roles or dozens of people may be in a single role. The main roles that are involved in the different stages described in the Assessment Life Cycle are shown below.

<b>Examination policy group</b>	<ul style="list-style-type: none"><li>• Makes critical decisions regarding exam policy</li><li>• Directors and Administrators of certifying organization</li></ul>
<b>Examination development group</b>	<ul style="list-style-type: none"><li>• SMEs, develops examination content and judges difficulty</li></ul>
<b>Practitioners</b>	<ul style="list-style-type: none"><li>• Active in profession, provides picture of professional practice</li></ul>
<b>Examination engineering</b>	<ul style="list-style-type: none"><li>• Psychometrician and related support staff</li></ul>
<b>Logistics personnel</b>	<ul style="list-style-type: none"><li>• Invigilators, test centre staff, candidate support</li></ul>

## EXAMINATION POLICY GROUP

The Examination policy group is generally composed of key stakeholders and decision makers along with a representative cross section of practitioners in the profession. For example, if the assessment program is a requirement that nursing students must pass into order to become licensed nurse, one would expect that the Examination policy group may be composed of:

- Executive director of the certifying organization
- Director of Operations of the certifying organization
- Project director of the examination program
- Representatives of the professional association (e.g., nurses, paramedics, pharmacists, etc.) that are subject matter experts. For example:

- o Instructors actively teaching in the profession (e.g., nursing professors)
- o Practicing members of the profession (e.g., licensed practical nurses with more than 10 years of experience)
- o Highly experienced members of the profession (e.g., licensed practical nurses with more than 25 years of experience)
- o Possibly a recent graduate of the program (e.g., new licensed practical nurse)

Generally, the Examination policy group would be between 5-30 people but there are no firm requirements on the size of the group. Because the group tends to meet in person there are cost considerations for keeping the group of manageable size but it is important to have all the right people in the same room together (whether that room is a physical conference meeting room or a virtual room).

The Examination engineering group (e.g., a psychometrician) would typically guide the Examination policy group.

This group provides the critical decision-making regarding the examination policy including:

- Establishment of the competency profile
- Determining blueprint weights
- Determining number of exam administrations
- Number of unique exam forms
- Length of exam
- Format of exam (computer-based or paper-and-pencil)
- Item development policy
- Validation and feedback of exam draft
- Locations where candidates will take exam and physical conditions of exam
- Administration & Invigilation protocol
- Documentation
- Security Policy
- Registration / Exam Granting
- Item quality / reliability policy
- Participate in standard setting decisions (e.g., cut score)

## EXAMINATION DEVELOPMENT GROUP

The examination development group is mainly composed of the Subject Matter Experts (SMEs) who have the knowledge, skills, and experience in the profession to be able to work from the competency profile and develop test questions for the examination. Depending on the size of the testing program this group could be as small as a handful of people to dozens of SMEs all dedicated to producing content for the examination program.

An Examination engineering group would generally conduct item writing training to this group in order to ensure that the standards for content development are well understood and the SMEs have a solid foundation of understanding for what makes a “good” test question.

Generally the Examination development group would be composed of:

- One or more “lead” authors that are the experts and train the less experienced content developers. These individuals may also take the role of reviewing questions that have been authored by other SMEs
- A significant number of professionals who are the subject matter experts and will be developing the content
- A professional editor is sometimes involved in this group to review all questions created to ensure that they meet the grammatical and structure guidelines of the examination program
- An item bank manager would take the role of coordinating item development efforts to ensure that all areas of the blueprint have sufficient questions authored ready for review

An important role of the SMEs in this group is to not only provide “raw” content but also to provide context for each question. For example, the SMEs should be providing difficulty estimates for each question that they author. This could be in the form of an estimated p-value (e.g., 0.600, 60% of candidates are expected to get this question correct) or in a more general format (Hard, Moderate, Easy). Also the SMEs should be providing information on which competency (e.g., specific, general, domain or area) the question is targeted to.

## PRACTITIONERS

The practitioner group provides the “real world” practical experience that serves as a reality check in the process. Sometimes member of the Practitioner group are also members of the Examination policy group to help ensure that the examination policy decisions are well grounded in a practical, applied foundation.

Though practitioners can serve in various groups, the purpose explicitly described in this white paper is to provide input into the nature of professional practice that can then be incorporated into the design of the exam. In particular, a representative group of them weigh in via survey on the prevalence and significance of competencies in the profile. Frequently, questions focus on various factors such as frequency (how frequently in a profession does an individual “do” the competency), importance (how important the competency is to the role of the professional), and / or criticality (how significant are the consequences if a practitioner makes an error performing the competency). This information is fed into weighting formulae directed by the Examination engineering group to arrive at a blueprint to be reviewed and finalized by the Examination policy group.

## EXAMINATION ENGINEERING

The Examination engineering group is the experts trained to analyze/interpret data and guide the process from a psychometric perspective to a defensible solution in all stages of the Assessment Life Cycle.

Generally the Examination engineering group is composed of psychometricians and their support staff.

The Examination engineering group is mainly responsible for:

- Calculating survey-based competency weights
- Generating draft blueprints based on input from the Examination policy group and Practitioner group
- Conducting item development workshops
- Conducting psychometric review of questions that are developed and field-tested
- Assembling the drafts of the examination form(s) based on constraints from stage 1 of the process
- Defining item development targets based on blueprints and form requirements
- Guide the Examination policy group
- Guide the Examination development group
- Conduct all psychometric data analysis, review, interpretation
- Conduct standard setting

## LOGISTICS PERSONNEL

The Examination logistics crew are the people on the front lines making the examination happen.

This group includes:

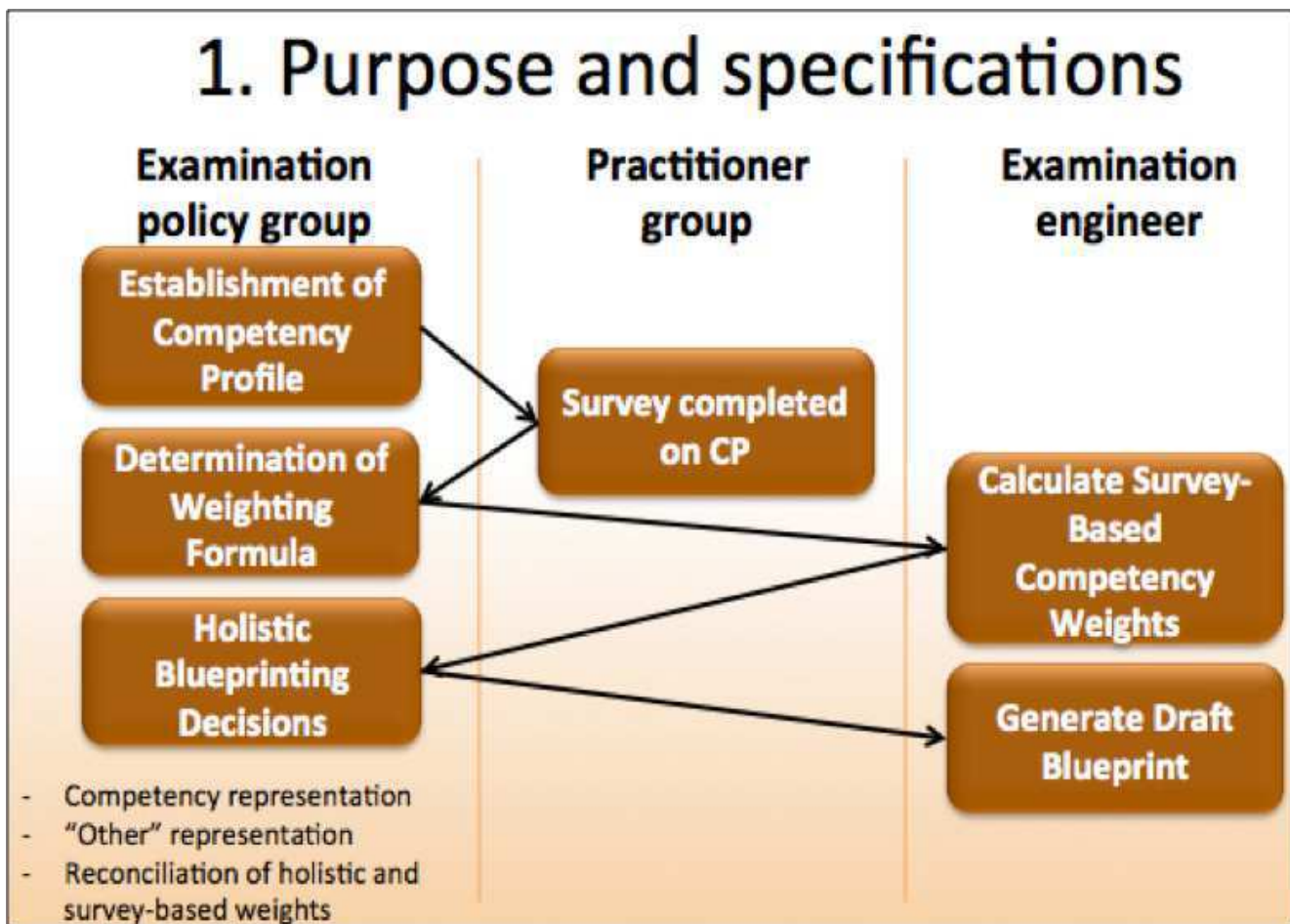
- Exam proctors
- Test centre staff and coordinators
- Candidate support

This group ensures that candidates have accounts created to take exams, that the test centre staff is prepared for the examination sitting, that proctors are trained and ready to invigilate the exams, and that standards are adhered to so that the candidates have the best experience possible to demonstrate what they know and can do.



## STAGE 1: DEFINING PURPOSE, TARGETS, AND CREATING SPECIFICATIONS

Stage 1 of the Assessment Life Cycle is where any assessment program starts. This stage begins by deciding what the assessments should be able to conclude about test takers, and finish by creating a clear blueprint on the number and types of test questions that are required to be written by the Examination development group.



In this stage, first the Examination policy group would convene to establish the competency profile. A competency profile is a set of behaviors that relate directly to the "on the job" work being performed. Often the competency profile includes levels of proficiency for each behavior. A competency profile can be anywhere from one to hundreds of pages depending on the profession and the levels of proficiency required/demonstrated in the profession.

Once a competency profile has been established the Practitioner group would be asked to take a survey to rate in various ways each competency. There are different approaches regarding what ratings the Practitioner group would provide in the survey. The importance, frequency, and criticality rating of each competency are common.

In the case where the Practitioner group has been asked to make multiple ratings on each competency, a decision must be taken to determine how the different ratings are to be combined in order to arrive at competency weightings. The Examination policy group (with the guidance of the Examination engineering group / psychometrician) will determine the weighting formula that generates the most satisfactory weights from the various ratings.

The most significant contribution of the Exam policy group to the final blueprints is the direct weighting of categories in the Competency Profile for the exam. Often, this process will proceed in two steps. First, the Group will rank order the categories according to their importance, then they will indicate the percentages of the exam that they judge ought to be assigned to each area. This is what is termed 'holistic blueprinting decision' in the figure above.

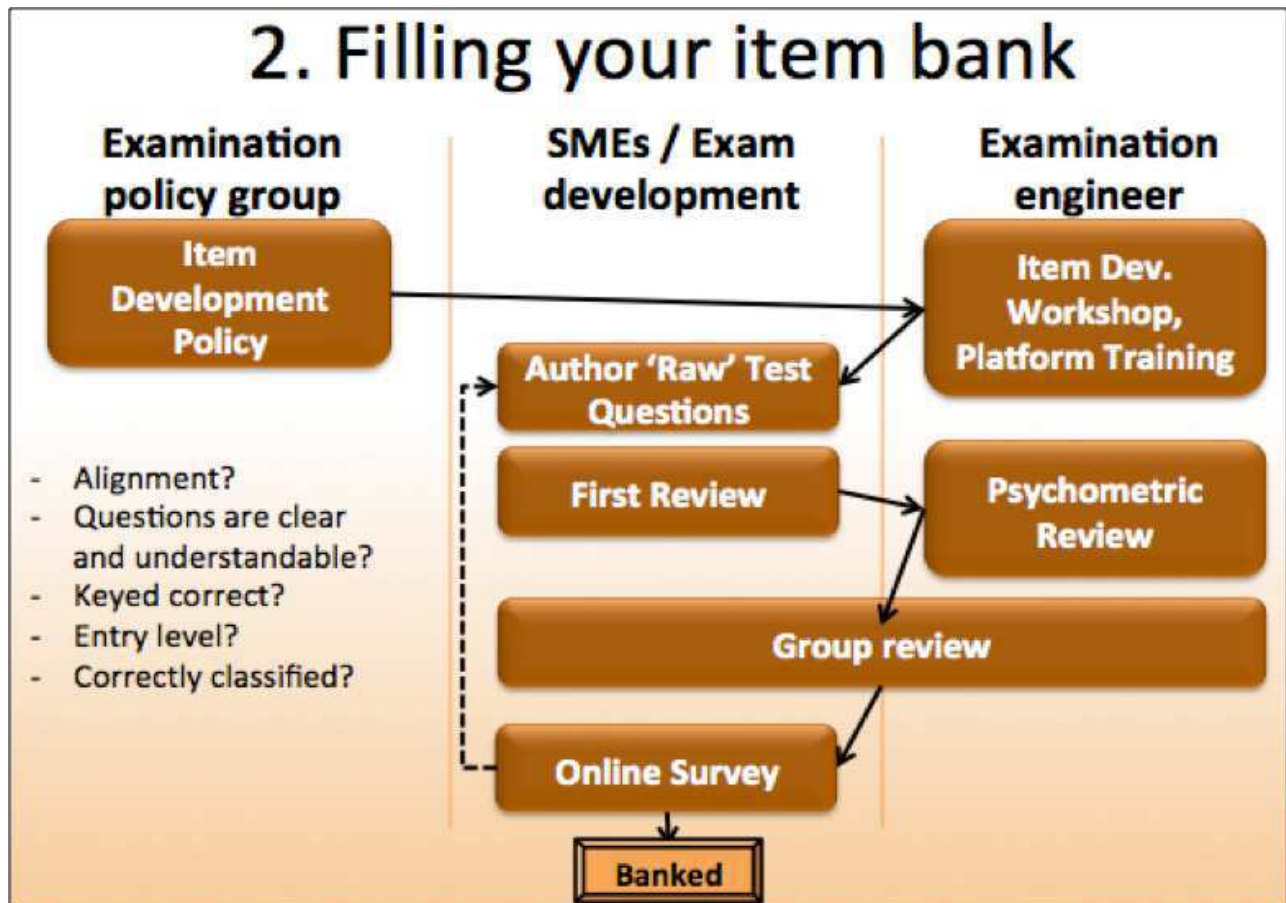
The next step is to compare the holistic decisions with the recommendations from the practitioner survey weighted by the formula determined earlier. When there are significant discrepancies, the Exam Policy Group may want to revisit their holistic judgments, at their discretion.

When this 'reconciliation' process between the holistic judgments and the practitioner survey are complete, a draft blueprint can be created.

With the exception of the establishment of the competency profile the other processes in this step can be accomplished in a relatively short period of time. Depending on the complexity of the competency profile, size of groups, and decisions that need to be made, this step can take as little as one day to several days.

## STAGE 2: FILLING YOUR ITEM BANK

Stage 2 of the Assessment Life Cycle is where the test questions are created to populate your item bank. In this stage the pre-specified number of items, based on the blueprint, number of forms, testing approach (e.g., fixed form, Computer Adaptive Testing – CAT, etc.) and other criteria of the testing program are written by knowledgeable SMEs to ensure that all the cells of your blueprint are nicely filled. The most important issue at this stage is to ensure that questions are constructed so they actually require the skill or knowledge indicated in the blueprint.



Based on initial guidance from the Examination policy group an item development policy document is drafted. This provides the foundation for the item development workshops and platform (software) training to be lead by the Examination engineering group. Although these sessions can be done virtually (i.e., online) many organizations opt to have the sessions hosted at an in-person physical location.

Once the SME and Exam development group have had thorough training on the platform in which questions will be authored, as well as training in the best practices surrounding item development, they will begin authoring questions. It can be the case that SMEs are tasked different sections of the blueprint to author (e.g., SME 1-5 tackle blueprint areas A-C, SME 6-10 tackle blueprint areas D-F, etc.) in order to most efficiently generate content to cover all areas of the blueprint.

Once “raw” questions have been created they go through a review process by senior/experienced SMEs and members of the Exam development group. A professional editor can also be involved at this point to review all questions. Generally this first review provides a comprehensive quality evaluation of the questions to ensure that item development standards surrounding blueprint alignment/classification, question clarity, grammar and wording standards are all adhered to. Any questions requiring edits can be sent back to the SME who authored the question for consideration to ensure that the intent of the question is retained and that the question can be improved appropriately. It is important that when creating questions that SMEs also assign a estimated difficulty level to each question. This could be in the form of a simple “Hard, Moderate, Easy” or a p-value type estimate (e.g., 0.600, 60% of candidates are estimated to get this question correct). These difficulty estimates can be used to help assemble the examination form(s) if field-testing the questions to obtain preliminary psychometric statistics is not possible.

Next a psychometric review of the questions occurs. The role of the Psychometrician or test development expert is to review the questions according to industry best practices as well as the standards articulate in the Item Development Policy document.

The last opportunity to modify the question comes in the group review. The purpose of the group review of each question is to determine which questions should:

1. Move forward “as is” to be added to an exam form
2. Require further edits to improve their quality and more correctly reflect their intent (i.e., re-field-test)
3. Be omitted completely from the item bank (i.e., not salvageable)

Group review can be a lengthy process so often this process has rules for reviewing questions. For example, no question will be reviewed for more than 10 minutes before a decision is made about what should happen to that question.

All questions that make it through the stringent review processes and psychometric scrutiny then undergo a final review process in which questions are presented to SMEs in a survey style format and a series of appropriate question are asked regarding each question. SMEs are asked to provide responses to survey questions for each candidate exam question regarding things like:

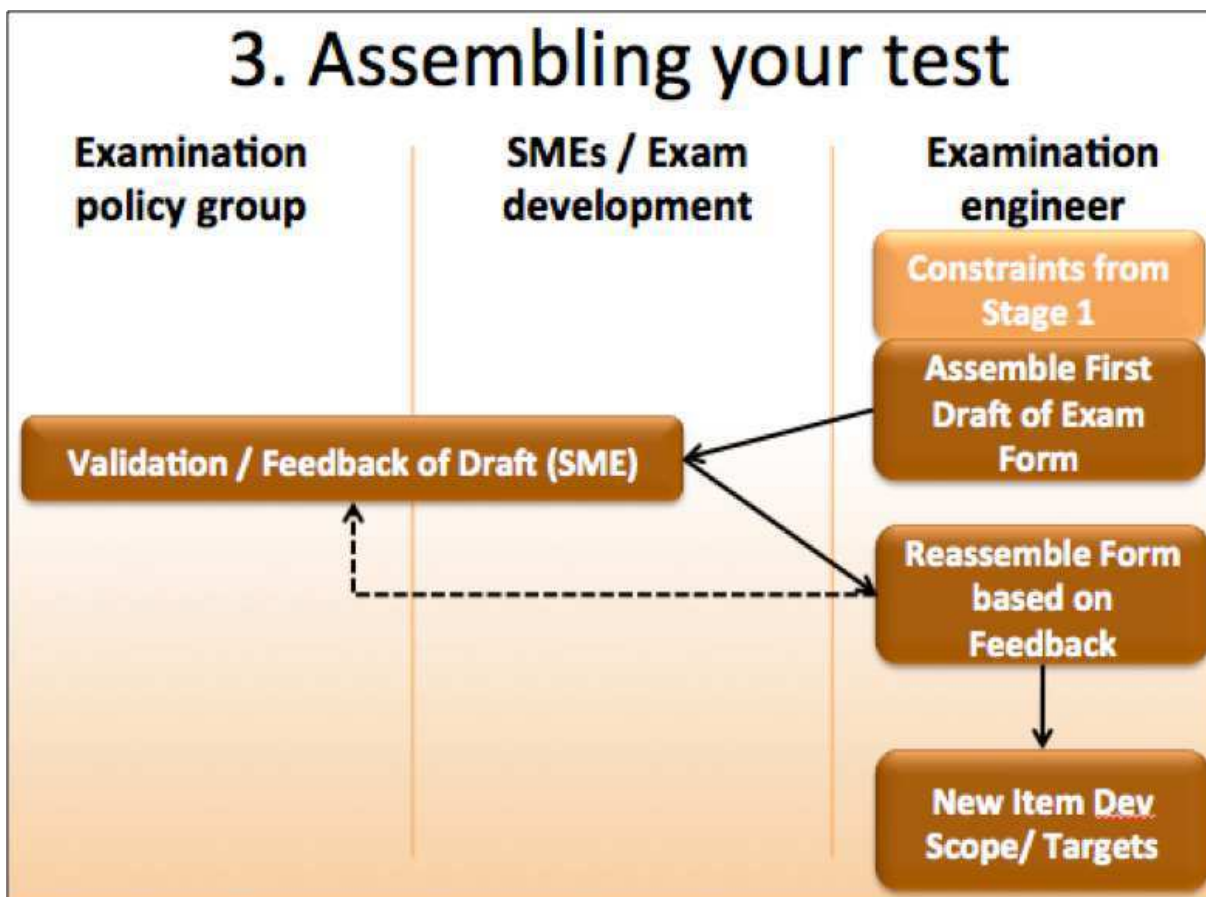
- Whether the keyed answer is appropriately, clearly, and accurately identified
- How strongly the SME agrees that the question is linked to the appropriate competency
- Their opinion of the accuracy of the difficulty estimate

Any questions that pass this final round of review are rightfully deposited in the item bank ready for field-testing. Often, these questions are added to operational test forms so that item statistics can be collected. If these statistics meet the standards in the Item Development

Policy, they may then be assembled into exam forms as operational questions and administered to candidates via the appropriate format.

### STAGE 3: ASSEMBLING YOUR TEST

Constructing your psychological measurement instrument (test/examination) is the essential part of Stage Three. Sometimes this is as easy as finding the right items from the right subject areas, but often, and particularly for certification/licensure exams, this step requires meticulous effort.



The first process in this step involves taking the constraints determined in step 1 (i.e., remember the blueprint?) and assembling a first draft of the exam form(s). Some examination programs may have more than one parallel form of the exam, which are each required to

adhere to the blueprint and also be equivalent to each other in terms of difficulty, reading length, etc. Often this process is done manually where an organization wades through the item bank pulling questions by hand that meet the appropriate criteria. As you can imagine, doing this by hand can be an extremely time consuming and iterative process. Fortunately, Automated Test Assembly (ATA) approaches using sophisticated algorithms makes this process more accurate and less time consuming.

Regardless of how the exam forms are assembled, the next step after preliminary assembly is to review the first draft of the exam form(s) by the Examination policy group and SME/Examination development group. The draft form(s) undergoes thorough review to ensure:

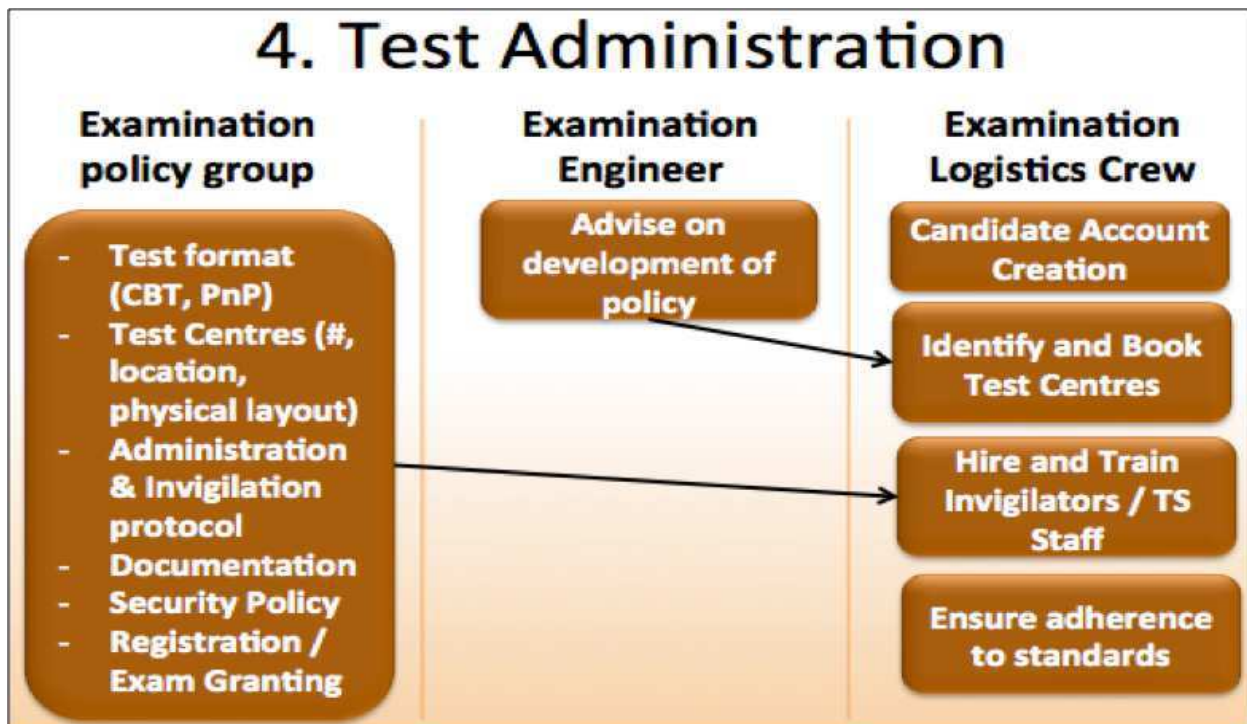
- Representation from blueprint is appropriate and accurate
- That no questions are “enemies” to one another appearing on the same form (e.g., “enemy” items are those that should not appear together as they may give away the answers to one another)
- That the “flow” of the exam is expected and appropriate
- That the exam form(s) are balanced in terms of question types, images, keyed correct answers, etc.

Based on the feedback obtained from the review a new draft version of the exam form(s) may be required. The form(s) are then reviewed again by the Examination policy group and SME/Examination development group until all stakeholders are satisfied with the quality and of the form(s).

Depending on how the review process progresses, new item development targets and scope may need to be added. For example, if in the exam form review process it is found that there is dramatic under-representation in one or more categories or aspects of professional practice, and the item bank has no more questions in those categories, then more item development will be required to bolster the number of items in those areas.

#### STAGE 4: TEST ADMINISTRATION

All the work that has gone into building the blueprint, creating the question content, establishing the item bank, and assembling the test forms is about to go to the next level. Stage 4 is where the rubber meets the road where candidates have the opportunity to demonstrate what they know and can do. The central concern here is to create conditions that are consistent with the ‘seriousness’ of the exam to ensure that candidates can successfully demonstrate their knowledge/skills/and ability via the measurement instrument.



In this stage the Examination policy group has a number of decisions that need to be in place for the test administration to go smoothly and appropriately given the examination program requirements. The Examination policy group will provide input on:

- Test format: Whether the exam is to be paper-based or computer based and if the later what format the exam should take (e.g., fixed form with multiple forms, CAT, etc.)
- Test centres: What number of test centres are required given the expected candidate volume and what area require test centres. Generally test centres are required within an approximate driving distance to a school/program (e.g., within a two hour drive of all dental hygiene schools in Canada) or some other criteria. Also, the Examination policy group determines the criteria for the exam in terms of test centre layout
- Administration & Invigilation protocol: Based on best practice guidelines and standards for high stakes examination administration the Examination policy group will determine the protocol that the tests will be administered under including the proctoring. Standards and guidelines such as those outlined by National College Testing Association (NCTA; <http://www.ncta-testing.org/resources/standards/index.php>) may be used as part of this decision making to guide this component

- Documentation: Producing templates and protocol for Irregularities reporting, In/Out reporting, proctor statements of understanding, etc are also generally crafted by the Examination policy group
- Security policy: Ensuring test security in all its forms is retained for an examination program is critical to its success. If candidates are cheating or stealing exam content the validity of the candidate scores are called into question undermining the defensibility of the testing program
- Registration/Exam granting: Who should be taking the exam (e.g., only candidates who are enrolled at an accredited institution with the permission of the department head) and how the listing of eligible candidates moves between systems is the responsibility of the Examination policy committee

The Examination engineering group provides review, input, and expert guidance based on industry experience. Any interruption of the candidate demonstrating what they know and can do will have validity implications and can throw the defensibility of the program into question. As such a great deal of careful thought and attention needs to be put into the documents and processes surrounding the exam.

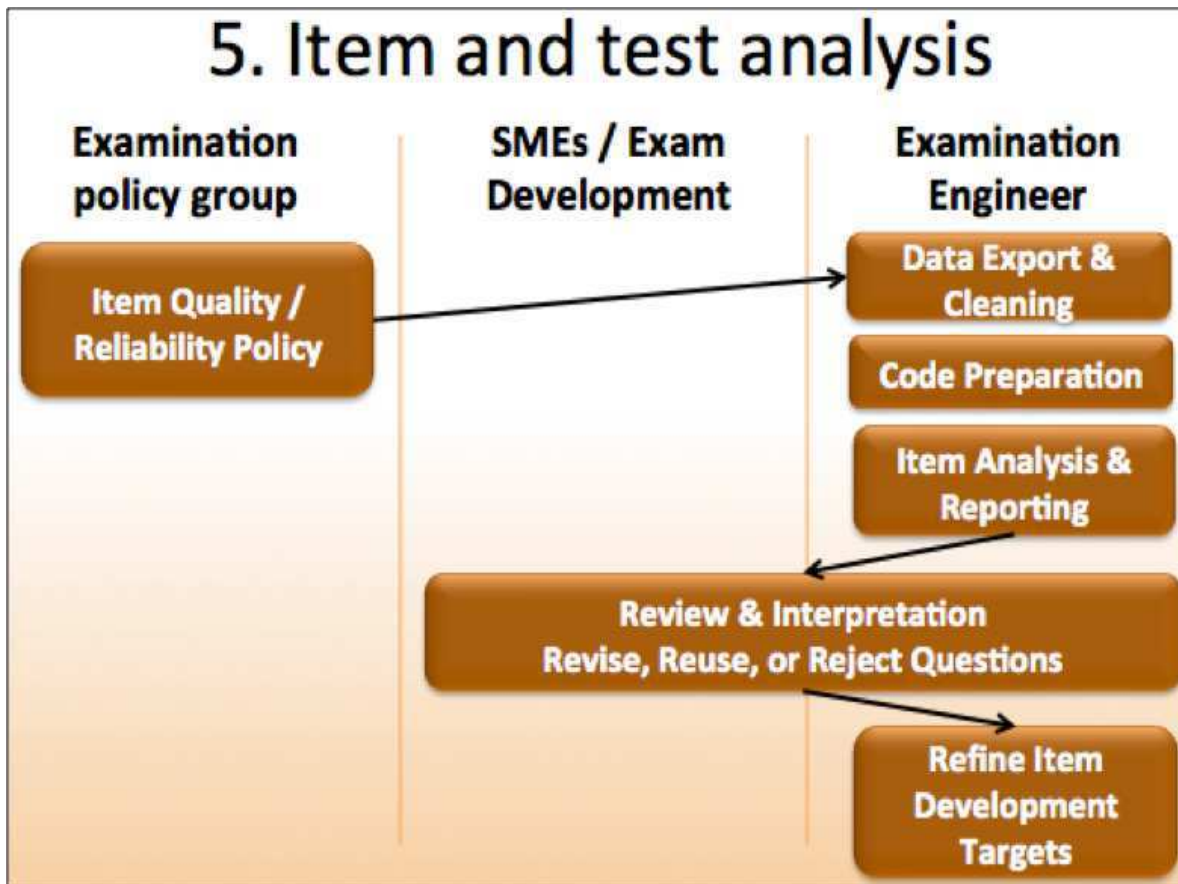
The Examination logistics crew is at the front lines of all the planning and implementation for an examination program. They ensure that the candidates have a place to take the exam and that all the logistics surrounding the defensible administration of the examination are in place and adhered to.

When everything is in place the candidates take their exams and in come the results to be hungrily analyzed by skilled psychometricians!

## STAGE 5: ITEM AND TEST ANALYSIS

Every Psychometrician dreams of having reams of good quality data coming in, ready to be analyzed and interpreted. In this step dreams become reality. In this stage, the actual data on how test takers responded to each item are analyzed to determine whether and how well the item is contributing the quality of the test.





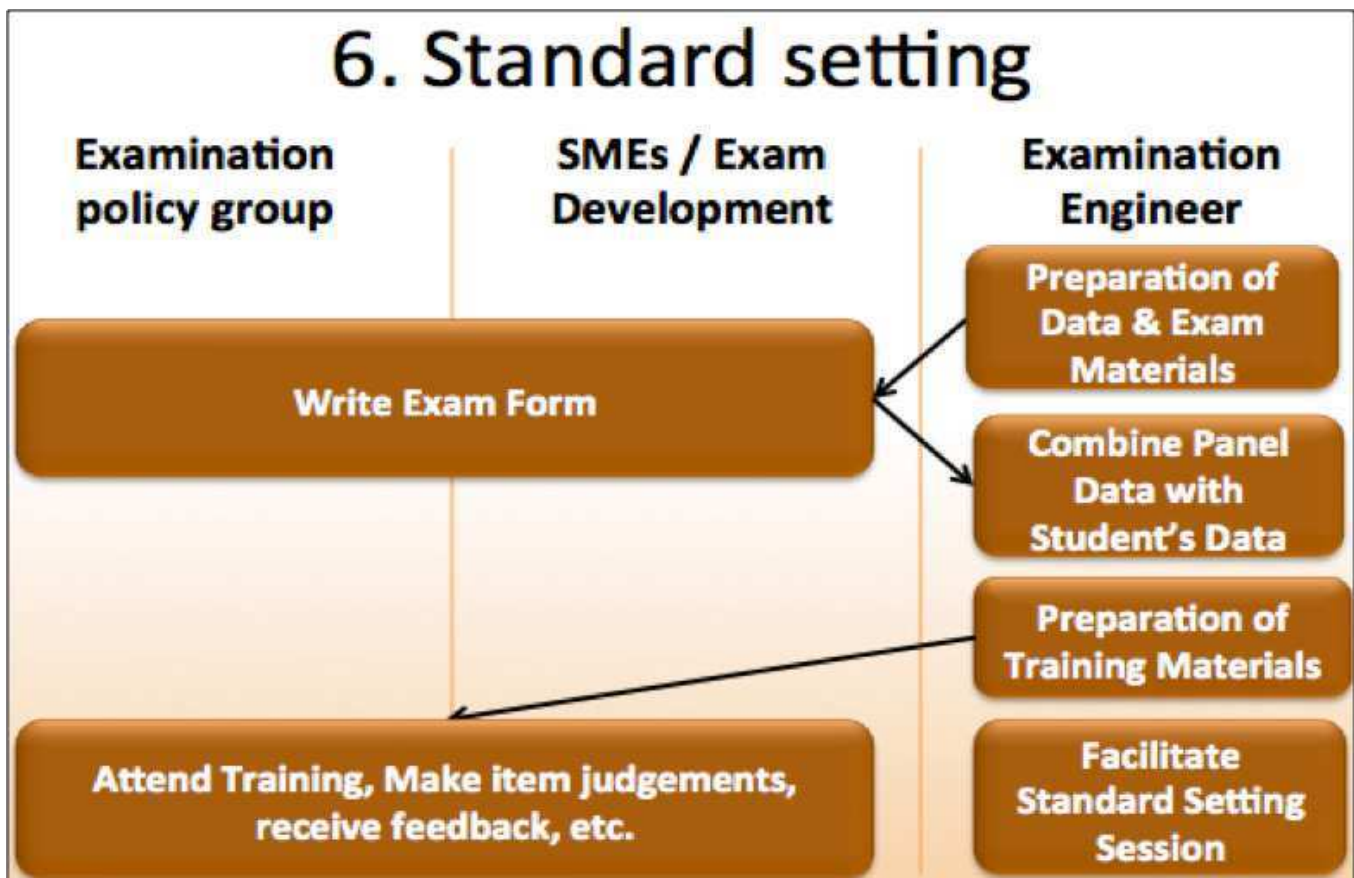
The first part of this process is for the Examination policy group to work with the Psychometrician to establish item quality guidelines and a reliability policy. The item quality guidelines could be very simple (e.g., p-values in the range of 0.2 to 0.9 and corrected item-total correlation coefficients must be greater than 0.300) or more detailed in terms of each item's performance. The reliability policy centres around what the targeted reliability of the assessment form should be. For example, for a high stakes certification exam the reliability policy may state that each exam form must have a Cronbach's Alpha internal consistency reliability of at least 0.9 or greater.

The rest of this step is very Examination engineering group heavy with the Psychometricians working their fingers to the bone to get all the test data analyzed, interpreted, and ready to communicate with the clients and stakeholders. Data from the examination session, such as candidate responses to each question, get formatted and run through item and test analysis programs to produce the psychometric information to be evaluated. The Psychometrician then reviews all the items on all the forms to ensure that they meet the requirements. This may

result in a refining of item development targets if a significant number of items are rejected or require further edits based on the thorough psychometric review.

### Stage 6: Standard setting

The last stage in the process is stage 6, standard setting. Standard setting is a critical step in any high stakes examination context where pass/fail decisions are being made. Essentially standard setting is the approach used to define proficiency, or achievement levels, and the cut-scores corresponding to those levels. The cut-score is simply the score that classifies the candidate into one level or another, for example the pass/fail cut-score on a certification exam would classify those candidate who achieve scores above the cut-score as passing (they know enough to be considered minimally competent) and those who achieve below the cut-score as failing (they do not know enough to be considered minimally competent). As you might expect, unless the cut-scores are appropriately, defensibly set, the decision to pass or fail any candidate could be called into question.



Generally standards setting sessions are done in person at a central location with all the relevant groups in the same room together. A standard setting session can take anywhere from 1-4 days with 2 days being the norm. The first process step in this stage involves the psychometrician getting all the materials prepared for a busy standard setting session. The Examination policy group and SME/Exam development group would then typically take the exam form as if they were a candidate and that data will be used for each member to get a clear perspective on the difficulty of the questions. This information is collated and added to the materials to be used as part of the session. During the standard setting session time is spent training the attendees on what is expected and how the process will unfold. Typical standard setting approaches are Modified Angoff and Nedelsky. To learn more about the specifics about each standard setting approach it is recommended to pick up a copy of the go-to book on standard setting, "Setting Performance Standards: Concepts, Methods, and Perspectives" by Gregory J Cizek. The Examination engineering group plays a pivotal role in facilitating the standard setting session and working with all groups to craft the appropriate cut-score(s) at the end of the session.

## CLOSING

The Assessment Life Cycle is a way of organizing the processes involved in creating sound assessments into a series of easy-to-understand stages. Following the Assessment Life Cycle will help ensure that your examination program is defensible and of the most benefit to candidates and your organization.

Let Yardstick help you apply the Assessment Life Cycle to your organization's assessment program. Yardstick offers a full range of products and services for every step and processes in the Assessment Life Cycle. Our clients agree, we know testing and we work hard to make your testing program the best that it can be.

Contact us at: [psychometrics@GetYardstick.com](mailto:psychometrics@GetYardstick.com) for more information, we would love to hear from you!

**LIST OF PSYCHOMETRIC SERVICES OFFERED IN EACH ASSESSMENT LIFE CYCLE STAGE**

**List of psychometric services offered in Assessment Life Cycle Stage 1: Defining Target and Creating Specifications**

Yardstick psychometric services division provides a host of services to our clients in each of the Assessment Life Cycle stages. Below is a list of psychometric services that Yardstick offers for

**Stage 1: Defining Target and Creating Specifications:**

Service	Description
Competency profile development	<p>The compilation of an inventory of skills and knowledge is a time consuming and detailed oriented process that is vital to the success of an assessment program. The competency profile (also referred to as Competency Models, Competency Profiles, Job Maps or Job Matrices) for a profession defines what professionals should know and can do within their job role. Often, a competency profile will not only specify which skills, knowledge, behaviors, or attitudes are required for mastery of the subject area, but also the importance and/or frequency of the competency areas.</p> <p>The Yardstick Psychometric Service division works with clients from start to finish developing competency profiles. We have experience in a number of domain areas including a range of health professions (e.g., opticians, nursing, physiotherapy, dental hygiene), science and engineering including the petroleum industry, and law enforcement.</p>
Blueprint development	<p>Once a competency profile has been developed an examination blueprint is required which outlines the specifications for an examination in terms of the importance weighting of each competency.</p> <p>The Yardstick Psychometric Service Division guides clients through the steps required in the development of a blueprint including:</p> <ol style="list-style-type: none"> <li>1. Surveying Practitioners</li> <li>2. Determination of a weighting formula</li> <li>3. Calculating Competency Weights</li> <li>4. Making holistic blueprint decisions</li> <li>5. Generating a draft blueprint</li> <li>6. Creating a finalized blueprint</li> </ol>

**LIST OF PSYCHOMETRIC SERVICES OFFERED IN EACH ASSESSMENT LIFE CYCLE STAGE**

Yardstick psychometric services division provides a host of services to our clients in each of the Assessment Life Cycle stages. Below is a list of psychometric services that Yardstick offers for

**Stage 2: Item development:**

Service	Description
Item development target planning	<p>Understanding how many items are required by competency domain and sub-domain area based on a blueprint and the requirements of an examination program is essential to building a robust item bank.</p> <p>The Yardstick Psychometric Service Division consults with clients to identify the scope of item development for a project (e.g., number of test forms, length of forms, common/unique item requirements for equating purposes, etc.), set item development targets, development item development project management strategies, etc.</p>
Item development workshops	<p>Item development is both a science and an art. Writing exam questions requires training and experience to produce high quality questions in a cost effective manner.</p> <p>The Yardstick Psychometric Service Division offers both face-to-face and remote item development-training workshops to ensure that item writers have the necessary knowledge and skills to produce the highest quality items possible to fill your item bank. Expertise is offered in the development of all question types including multiple choice, multiple response, and constructed response.</p>

<p>Performance based task development</p>	<p>Many organizations especially in the health sciences domain require practical examinations in the form of Observed Structured Clinical Examinations (OSCE). Development of these tasks, and the defensible scoring models surrounding these tasks, require skill and expertise to master.</p> <p>The Yardstick Psychometric Service Division offers professional performance based task development in all testing domains including pilot testing and evaluation of new OSCE tasks.</p>
<p>Professional item editing and review</p>	<p>After raw items are produced these items need to be edited and reviewed by individuals skilled in professional item development.</p> <p>The Yardstick Psychometric Service Division offers professional item editing and item review services in which items are scrutinized to ensure grammar, punctuation, structure, formatting, and content domain appropriateness are of the highest quality for candidates.</p>
<p>Professional item development</p>	<p>Finding Subject Matter Experts that can spare the time required to create large numbers of items can be challenging. As such the Yardstick Psychometric Service Division offers professional item development services including item cloning and Automated Item Generation to help bolster client's item banks without sacrificing the defensibility of the item development and review process.</p>
<p>Translation services</p>	<p>Providing translations of questions in multiple languages for candidates is a necessity for many of Yardstick's clients. As such the Yardstick Psychometric Service Division offer professional translation services to take items in a source language and translate them successfully into one or more target languages as required.</p>

### LIST OF PSYCHOMETRIC SERVICES OFFERED IN ASSESSMENT LIFE CYCLE STAGE 3: ASSEMBLING EXAM FORM(S)

Yardstick psychometric services division provides a host of services to our clients in each of the Assessment Life Cycle stages. Below is a list of psychometric services that Yardstick offers for

#### Stage 3: Assembling exam form(s):

Service	Description
Automated Test Assembly (ATA)	<p>Assembling an exam form manually by searching through an item bank of any reasonable size and piecing together an exam form item by item is time consuming and monotonous work. If the task is to assemble dozens of exam forms with 200 items composing each and a list of blueprint and other constraints the manual task become monstrous and non-scalable.</p> <p>To solve this problem for our clients the Yardstick Psychometric Service Division offers professional Automated Test Assembly (ATA) services which us advanced solver algorithms to automatically assemble exam forms in a fraction of the time of manual form assembly and with much great accuracy.</p>
Exam form validation facilitation	<p>Once an organization has the exam forms assembled validation of those forms needs to occur. Generally Exam Managers, Subject Matter Experts, and other roles will be involved in the validation of examination forms.</p> <p>The Yardstick Psychometric Service Division offers to facilitate the exam form validation process to ensure that all exam forms preceding an administration are thoroughly reviewed and judged to be defensible.</p>

## LIST OF PSYCHOMETRIC SERVICES OFFERED IN ASSESSMENT LIFE CYCLE STAGE 4: TEST ADMINISTRATION

Yardstick psychometric services division provides a host of services to our clients in each of the Assessment Life Cycle stages. Below is a list of psychometric services that Yardstick offers for

### Stage 4: Test administration:

Service	Description
Paper-based to computer-based transition planning	<p>Administering examinations on paper is a tried and true method but computer based exam administration offers many advantages including rapid results reporting, less printing, shipping, and storage of paper materials, and innovative interactive assessment formats not available in a paper format.</p> <p>Migrating clients from paper-based administration of examinations to computer-based administrations is a specialty of Yardstick's. We have successfully migrated dozens of organizations from their legacy paper-based administration processes to leading edge computer-based exam administration processes.</p>
Collusion detection and test security analyses	<p>Post-administration there can be speculation that cheating or content theft had occurred during an exam administration. Fortunately there are statistical methods to detect whether cheating has occurred during an exam administration at specific test centres right down to specific candidates who may have been involved.</p> <p>Yardstick offers a host of collusion detection and test security analyses for our clients to shed light on potential security issues.</p>
Bubble sheets design and result analysis	Should paper-based exam administration be a necessity Yardstick offers bubble sheet design and analysis services to make the scanning and data processing of bubble sheets straightforward.
Customized candidate results reports	Whether reports are required for all candidates or just for those that failed an examination in order to provide diagnostic feedback Yardstick can help. We have provided organizations with many options for customized feedback reports to candidates, administrators, and other stakeholders.



## LIST OF PSYCHOMETRIC SERVICES OFFERED IN ASSESSMENT LIFE CYCLE STAGE 5: ITEM AND TEST ANALYSIS

Yardstick psychometric services division provides a host of services to our clients in each of the Assessment Life Cycle stages. Below is a list of psychometric services that Yardstick offers for

### Stage 5: Item and test analysis:

Service	Description
Classical Test Theory (CTT) item analysis	<p>Classical Test Theory (CTT) item and test analyses are by far the most common type of psychometric analyses performed on questions and exams in the assessment industry. CTT provides highly useful diagnostic information regarding the performance of items and exams even in conditions where sample sizes are an issue.</p> <p>Yardstick provides tailored CTT item analysis reports to organizations, which diagnose the psychometric performance of items. These reports and interpretation services are crucial to the defensibility of an examination program and provide valuable feedback to organizations regarding areas where item development can be improved.</p>
Item Response Theory (IRT) item analysis	<p>Item Response Theory (IRT) provides rich information regarding items with different assumptions than CTT making it useful for a range of applications such as Computerized Adaptive Testing (CAT). Clients with larger sample sizes can take advantage of IRT to evaluate the performance of their item bank.</p> <p>Yardstick has extensive in-house expertise in the area of IRT analysis and application including CAT, Differential Item Functioning analysis, and collusion detection approaches that utilize IRT.</p>
Differential Item Functioning (DIF) analysis	<p>Differential Item Functioning (DIF) analysis allows for the investigation of how items perform for different groups of candidates. For example, English and French speaking candidates to take the same set of questions. Ensuring that items perform fairly for all groups of candidates is essential for assessment program defensibility. Yardstick offers DIF analysis expertise to our clients.</p>

**LIST OF PSYCHOMETRIC SERVICES OFFERED IN ASSESSMENT LIFE CYCLE STAGE 6: STANDARD SETTING**

Yardstick psychometric services division provides a host of services to our clients in each of the Assessment Life Cycle stages. Below is a list of psychometric services that Yardstick offers for

**Stage 6: Standard setting:**

Service	Description
Standard setting facilitation	<p>Standard Setting, the business of setting passing scores, (AKA ‘cut scores’), for examinations is an important stage in the Assessment Life Cycle. The cut score(s) serves to classify candidates into categories; those who score above the cut score are judged to possess the minimum level of proficiency required for inclusion into the category, while those who score below the cut score are deemed not to possess the minimum level of proficiency required, and are classified accordingly.</p> <p>Standard Setting is most often a judgmental process. It typically involves a panel of subject matter experts and stakeholders who must estimate the difficulty of each question for so-called minimally-competent, or borderline candidates. Those judgments are then aggregated in order to arrive at a passing score across sets of questions. Standard setting methods differ in the way that those judgments are made by panelists, and also in the way that question-level judgments are aggregated to create a passing score. Three common methods of standard setting are the Modified Angoff method, the Bookmark procedure, and the Nedelsky method.</p> <p>Yardstick has a wealth of experience conducting standard setting sessions around the globe in numerous organizational contexts.</p>
Automated Angoff estimation services	<p>With very large items banks it can be financially challenging to standard set every item. For this reason Yardstick offers Automated Angoff estimation services, which use leading edge algorithms to estimate Angoff values for items based on characteristics and data associated with each item in the bank.</p>