STAGE 2

Item Development

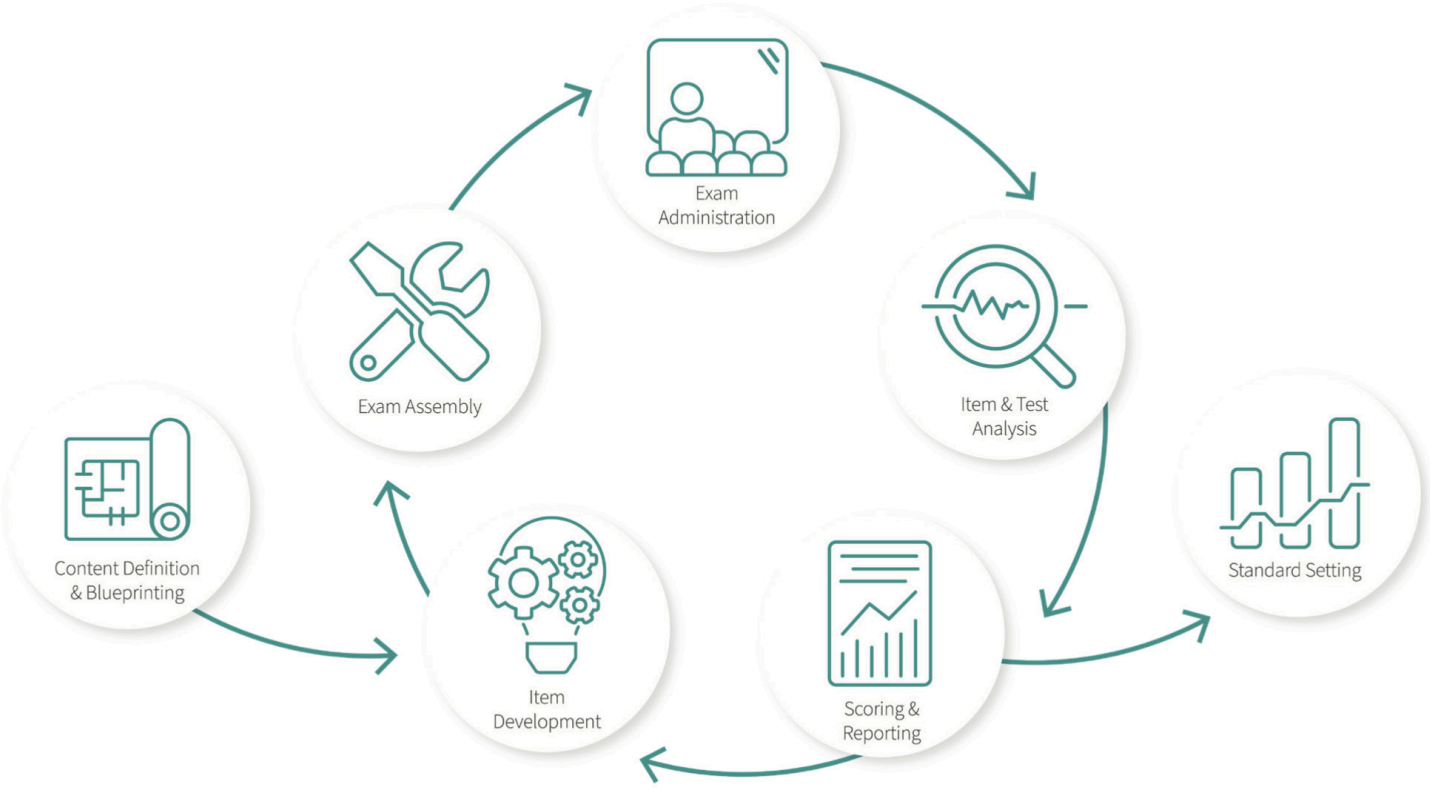
MEAZURE LEARNING

Meazure Learning Whitepaper

Table of Contents

Introduction

The Assessment Life Cycle is a way of organizing the processes involved in creating valid assessments into a series of easy-to-understand, logical stages.



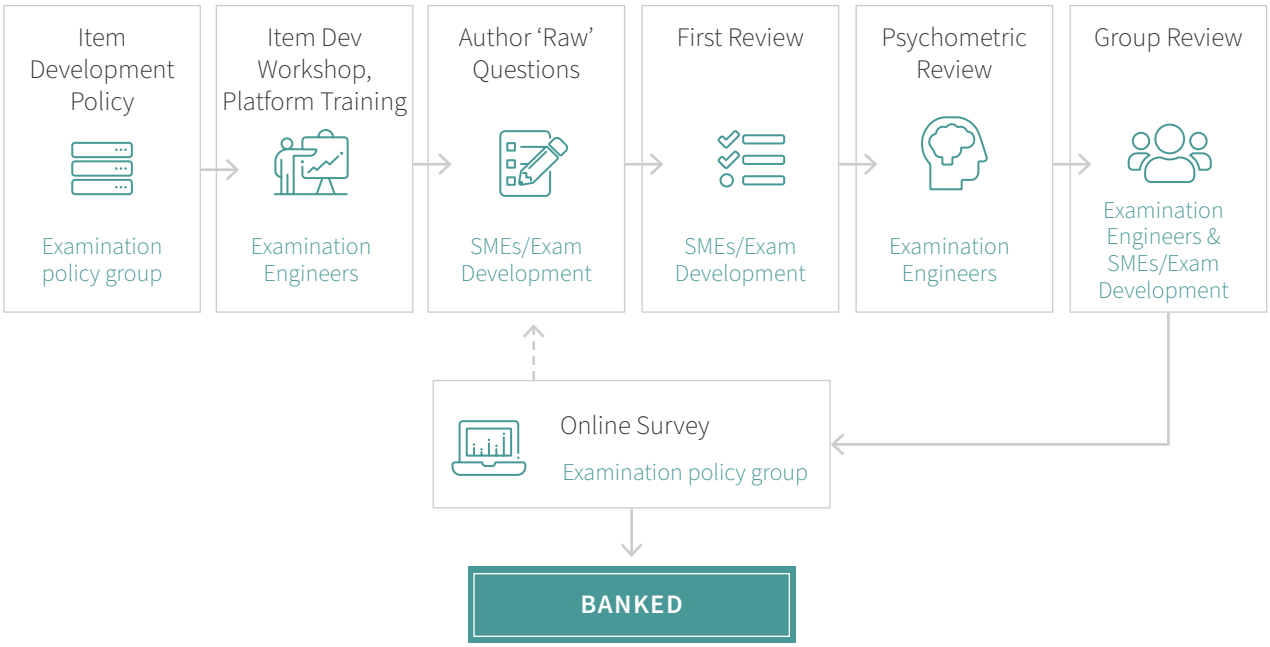
Stage 2 of the Assessment Life Cycle is where test questions are created and used to populate an item bank. That may sound simple enough. But in truth, item development is careful work. It requires the coordinated effort of several groups of experts to ensure that the final examination includes the right number of items, organized into the right categories.

Several parties are involved at this stage of the examination development process. First, there's the examination policy group, who determines how the item development, review, and validation processes should be designed. Next, there's the examination development engineer (or psychometrician), who conducts item development training, supports the item writing process, and facilitates a review of

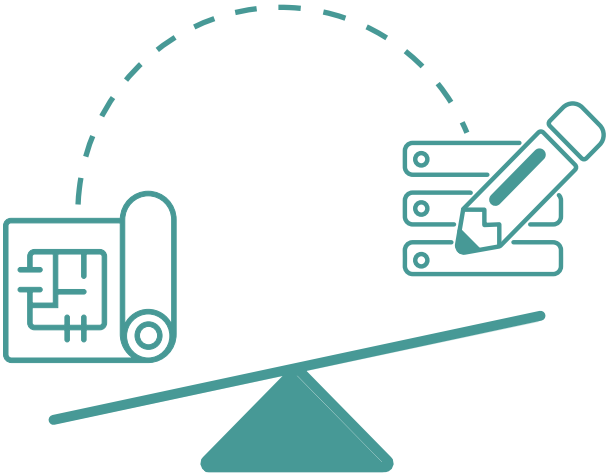
the items that are written. Finally, there are the subject matter experts (SMEs), who author the actual examination questions. It is important to carefully coordinate the work of these three parties in order to ensure that the items being written are a valid representation of the specifications in the examination blueprint (developed back in Stage 1). Ensuring your item bank has the right amount and right kinds of questions will go a long way towards making sure that your examination fits the unique requirements and constraints of your certification program at large.

To get a sense of the specific tasks that occur during Stage 2 of the Assessment Life Cycle, consider the figure below:

FILLING YOUR ITEM BANK



Throughout the item development process, one key concern is whether the item bank is achieving something called "alignment". Alignment is the extent to which the actual examination matches its corresponding foundational test documents (such as the competency profile and examination blueprint). Remember that an examination blueprint in particular is meant to serve a 'road map' for item development. It specifies the desired number of questions that a given examination should contain for each competency and competency area (in addition to other specifications such as cognitive level, population type, and number of case questions versus individual questions). When an examination meets these specifications, it can be said that the items in the examination are aligned with the blueprint. In other words, the examination displays evidence of content validity.



This is no easy feat. In order to satisfy all of the design constraints, many different characteristics of the test questions must be carefully engineered and balanced. The first set of constraints are those outlined in the examination blueprint. These include content constraints (such as the topic, unit, or competency area being assessed), the level of thinking required for answering questions (e.g., knowledge/comprehension, application, or critical thinking), and the format of the questions themselves (e.g., multiple-choice, short answer, or drag-and-drop).

The second set of constraints is meant to minimize bias in the examination. These constraints aren't always the responsibility of the item writer per se; but often come into play when the test is being assembled from the available items in the item bank. Given a robust item bank, there are often many potential examinations that could match the blueprint specifications. That said, not all versions of the examination will be equally unbiased. Some potential examinations can inadvertently advantage some candidates over others for reasons that are unrelated to candidates' actual ability in the subject area. To minimize bias, certain considerations must be taken into account, including: the proportion of different correct answers on the multiple-choice answer key (i.e., 'A', 'B', 'C', or 'D'); the number questions featuring different populations of interest (e.g., male versus female); and, the number of "case-based" questions that require candidates to read and connect a written passage to a series of related questions.

Often, there is a specific individual (i.e., the item bank manager) who takes responsibility for

facilitating and overseeing the item development process. This individual is tasked with making sure that both types of constraints listed above are properly accounted for. As part of this role, the item bank manager ensures that each question proceeds through something called an "item development workflow" – a sort of quality control process. This process is meant to ensure that every question that makes it to the final item bank has been thoroughly vetted in terms of its correctness, alignment, and defensibility; and that each meets best practice psychometric standards for high-quality examination questions.

Now that we know what the goals of item development are, let's consider the process that should be followed to make item development work - and work well.

1 Drafting an Item Development Policy Document

The first step in the item development process involves the creation of an item development policy document. This document is usually jointly drafted by the examination policy group and the examination engineer/psychometrician.

The purpose of this document is to clearly lay out the roles and responsibilities of each of the groups involved in the item development process. Although the length and specificity of this document may vary between clients and examinations, what is important is that it contains a record of all of the decisions related to requirements gathering and stakeholder meetings. This document usually includes information on:

- The establishment of the competency profile
- The determination of examination blueprint weights
- The number of examination administrations
- The number of unique examination forms
- The length of the examination
- The format of the examination (i.e., is it computer-based or paper-and-pencil?)
- The item development policy
- The validation and feedback of examination drafts
- The location(s) where candidates will take examination; and the physical conditions of examination
- Protocols related to the administration and invigilation of the examination
- The process for documentation
- The examination security policy
- Information on registration and examination granting
- The item quality/reliability policy
- Any decisions related to the establishment of an examination cut score (i.e., standard setting)

As you can see, the item development policy document is fairly comprehensive; and by enshrining all of this information, it can effectively guide the development of the item bank and examination configuration; as well as the logistics for later stages in the Assessment Life Cycle (for example, determining how many test centre locations are required, and in what locations).

2 Item Development Workshop(s) and Platform Training

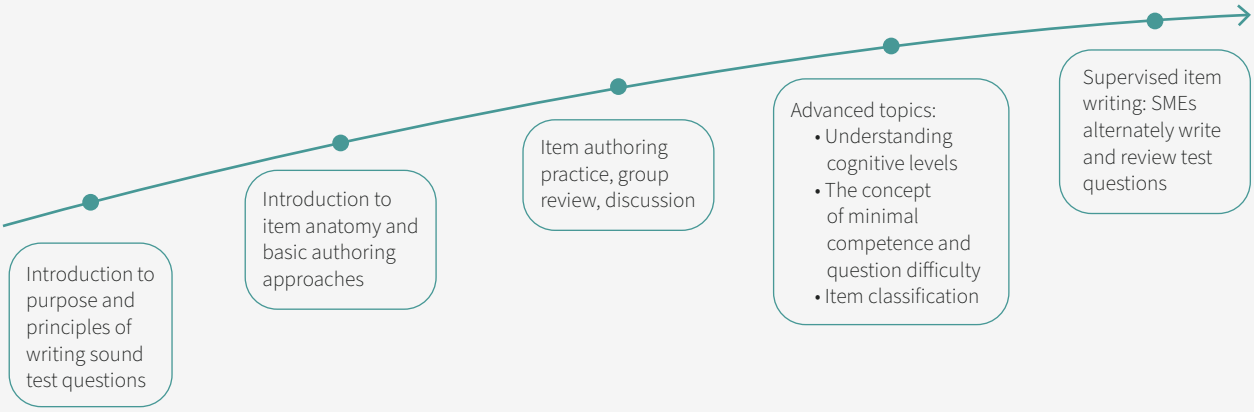
Once the examination policy document has been written and approved by all parties, the next step is to train the subject matter experts (i.e., item writers) on the process for writing questions using the selected item banking platform. To ensure the rest of the item development process runs smoothly, it is important to ensure that all subject matter experts have sufficient training on item development best practices, as well as how to navigate the item banking platform.

Although these training sessions can be held virtually, many organizations choose to have them hosted at an in-person physical location instead. Although this is more costly, in-person sessions not only provide the raw information regarding item development best practices; but also provide important professional development opportunities where the subject matter experts can form professional relationships with one another, the members of the examination policy group, and the examination engineering group.

ITEM DEVELOPMENT WORKSHOP(S)

The idea behind an item development workshop is to provide the prospective item writers with the knowledge and skills they need to write high-quality questions that meet psychometric best practice standards. These workshops are usually run by a psychometrician or other test development specialist and may last anywhere between half a day to five full days.

To give you an idea of the process of running an item development workshop, consider the example below. Although this is a good overview of the steps involved, know that some details will vary depending on the number of attendees and the length of the workshop (for example, some processes listed below may occur more than once):



It is not unusual for some subject matter experts to take to item development more quickly (or with more enthusiasm) than their peers. When facilitating such a session, it is important to watch for this, do what one can to engage the other subject matter experts, and (potentially) explore the possibility of promoting the individual to a more senior item development role.

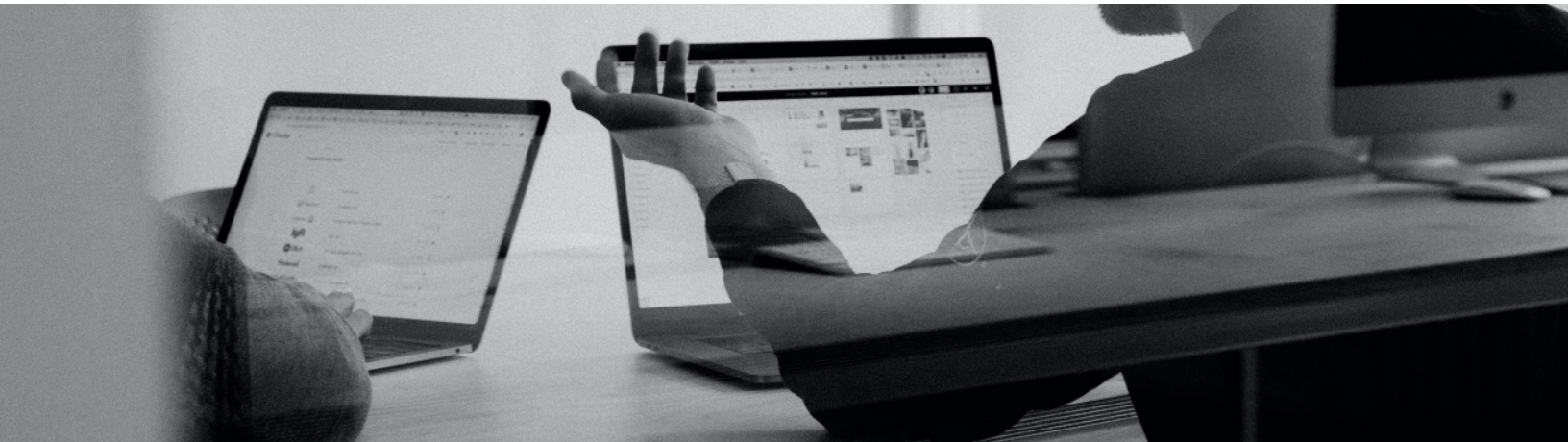
PLATFORM TRAINING

Access to item banking software is often strictly managed. Consequently, authoring questions using an online platform can significantly promote item and test security (making online item banks an invaluable tool for item development).

In order to streamline the training process, it is often a good idea to include an orientation to the item banking platform as part of the item development workshop. This step should not be overlooked. It is important to ensure that the subject matter experts know how to enter the questions they create into the item banking system. Making sure they can do this effectively will help ensure that there are no barriers to producing a significant volume of high-quality questions. In addition, many item banking systems (e.g., Measure, which Meazure Learning uses) categorize items in the bank using codes that the item writers assign. Making sure that your subject matter experts know how to navigate the item banking system can also help ensure they enter in the correct supplementary codes or other information when creating questions (ensuring the system categorizes the question appropriately).

Although important, platform training does not have to be an onerous process. If the software application is intuitive, a brief overview of the system (with some hands-on exercises) will get a keen group of item writers off to the races! This basic training should focus on:

1. How to write questions and indicate the correct answer using the software provided
2. How to edit questions and submit them for further review
3. How to fill in supplementary ‘metadata’ (e.g., competency, level of knowledge, etc.)
4. How to write passages and link them to questions, in order to create case-based questions



3 Implementing and Managing the Item Development Workflow

Once the subject matter experts have been trained on how to write high-quality questions and enter them into the item banking platform, they will begin authoring their own examination questions. To help ensure that all areas of the examination blueprint are covered, different subject matter experts may be tasked with writing on different sections of the blueprint. This is a good approach for efficiently generating content that will cover all areas of the blueprint; and can leverage the fact that different subject matter experts will have different areas of expertise (and may be more comfortable writing questions for some competency areas rather than others).

Once "raw" questions have been created, they must go through a review process prior to being approved for inclusion on the final examination. There may be several iterations of item review, including a psychometric review (focused on ensuring the item follows item writing best practices and the examination policy development document), a content review (focused on ensuring the question adheres to the examination blueprint, tests candidates at the correct level, and has only one correct answer), and an editorial review (focused on ensuring items have correct spelling, grammar, and avoid jargon or unclear phrasing).

These reviews normally involve a psychometrician or other test development expert, as well as a team of senior or experienced subject matter experts and members of the examination development group. A professional editor can also be involved at this point to review all questions.

Reviewing item content requires a great deal of subject matter expertise. Consequently, it is often helpful to hold a group review with subject matter experts (e.g., the group of item writers, or a separate examination review group). The purpose of the group review is to determine which questions should:

1. Move forward "as is", and be added to an examination form.
2. Be edited further to improve their quality and more correctly reflect their intent.
3. Be omitted completely from the item bank.

Group reviews can be a lengthy process. To keep the group on track, it is often helpful to have 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.

Any questions that require editing after they are reviewed can be modified by the item reviewer(s). Alternatively, they can also be sent back to the original question writer for revision. The latter approach is often advisable, as having the input of the original author ensures that the intent of the question is retained when reworking it.

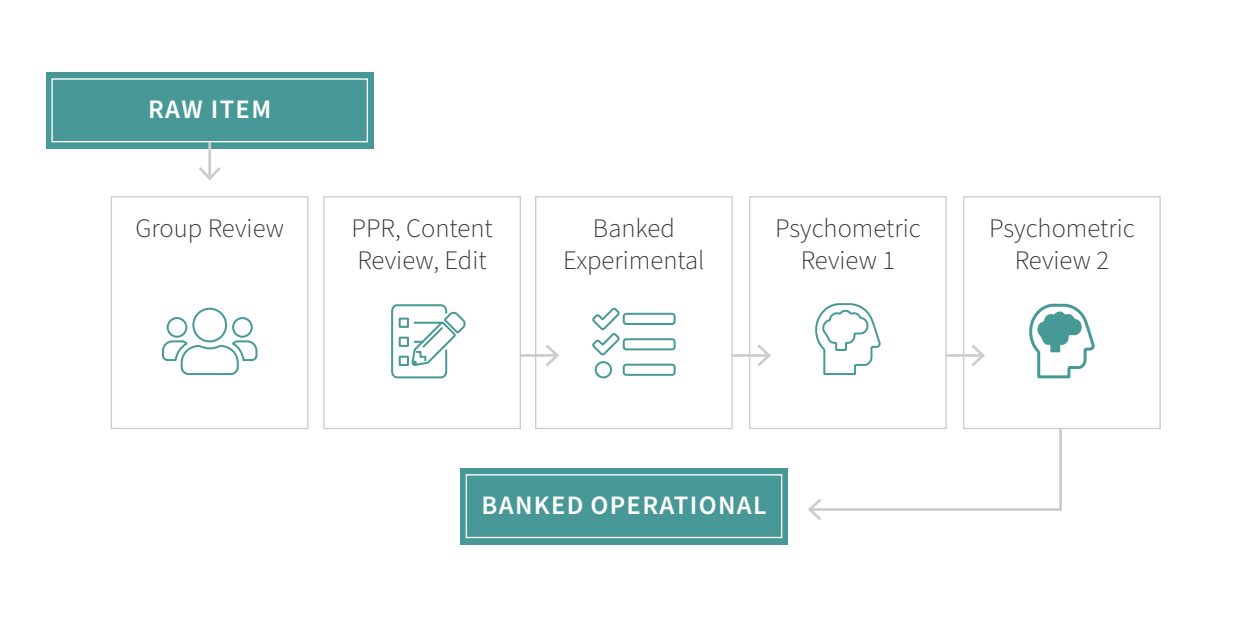
After questions make it through this stringent review process, they often undergo one final review. At this stage, each question is presented to a group of subject matter experts; and a series of final questions is asked regarding each question. For example:

- Is the keyed answer clearly correct (or are other alternatives possible)?
- How strongly is the question linked to the appropriate competency?

Any question that passes this final check is then deposited into the item bank, ready for field-testing. To field test a question, it is often added to an operational examination as a non-scored item. This way, statistics on the question's difficulty, its ability to discriminate between candidates, and the quality of its distractor options can be collected without affecting the score of candidates taking the examination.

If these statistics meet the standards set forth in the policy guide, then the question may be included in future examination forms as an operational, scored item.

To give you a visual representation of the process, consider the figure below:



Now that we've considered the workflow, let's examine the specific roles that individual subject matter experts and test development specialists play in making it all happen.

Item writer(s):

- Obtains guidance from the item bank manager/psychometrician regarding which competency areas they should write questions for.
- Writes the initial test questions, and classifies them along various categories (e.g., competency, cognitive level) using 'metadata'.

Professional practice reviewer(s) (validation):

- Has a sound 'big picture' view of the profession and the item development process. Uses this perspective to determine whether questions can be entered into the item development workflow.
- May also engage in item content review, or editing test questions.

Content reviewer(s)/editor(s):

- Reviews questions and either edits them or sends them back to the original author, so that they can be improved and rendered consistent with professional practice standards.
- Ensures that items are written clearly, so that test-takers will understand what is being asked of them.

Psychometrician(s)/test development specialists(s):

- Gives item development and platform training; and facilitates item writing and review sessions.
- Reviews questions to ensure that they are consistent with psychometric best practices.

Team leader for group review:

- Serves as a subject matter expert with a sound 'big picture' view of both the profession and the item development process.
- Facilitates decision-making regarding potential final adjustments to test questions.

Item sign-off/validator:

- A group of subject matter experts who review each question via survey/checklist, in order to ensure each question's compliance with a set of pre-established quality criteria.
- Give items a final "stamp of approval".

Item bank manager:

- Serves as the project manager for the item bank development process.
- Ensures that questions are created efficiently and according to targets set forth in the examination blueprint.

As you can imagine, the process of creating a large number of high-quality test questions that cover the breadth of the blueprint is a formidable task. It takes the coordinated effort of many qualified and motivated individuals to make it all happen. However, when this process is managed well, your new questions will effectively assess test-takers on the subject area; and will do their part to support the defensibility of your assessment program. Once you have a robust bank of high-quality items, it should be easy to move on to the next stage of the Assessment Life Cycle (described in brief below).

Next Stage | Assembling your Test

Once you have a robust item bank, the next stage in the Assessment Life Cycle involves drawing items from the bank and creating your examination form(s). To do this, you need to review the examination constraints determined in Stage 1 (i.e., those in the examination blueprint), and assemble a first draft of the examination form (or forms).

Some examination programs may have more than one parallel form of the examination. If this is the case, then each form of the examination should adhere to the examination blueprint. Moreover, each examination form should be equivalent in terms of its overall difficulty.

Often, this equivalence is achieved by manually searching through the item bank and pulling questions that meet the appropriate criteria. This can prove a time-consuming and iterative process. Fortunately, there is an alternative called "Automated Test Assembly" (ATA), which uses sophisticated algorithms to compute the best possible combination of items to meet your blueprint specifications. This makes the test assembly process more accurate (and much less time-consuming). We will discuss how this process is done (and much more) in the next white paper in our Assessment Life Cycle series!

Conclusion

In summary, the Assessment Life Cycle is a way of organizing the processes involved in creating valid assessments into a series of easy-to-understand, logical stages. The focus of this whitepaper was to detail the fundamental steps and key processes that are involved in the second of these stages (i.e., item development). As covered, this involves a number of specific steps, including:

1. Establishing an item development policy (and policy documentation)
2. Facilitating item development workshops and platform training
3. Writing and monitoring the creation of new items
4. Reviewing newly-created items and advancing them through the item development workflow

Following these best practice steps – and the Assessment Life Cycle in general – will help ensure that your assessment program is valid and defensible; affording the greatest possible benefit to both your test-takers and your organization.

Let Meazure Learning help you apply the Assessment Life Cycle to your assessment program. Meazure Learning offers a full range of products and services that cover every step and process. Our clients agree: we know testing; and we will work hard to make sure that your testing program is the best that it can be.

To explore this opportunity – or for more information – please feel free to visit us at:

meazurelearning.com/services

List of Psychometric Services offered in Assessment Life Cycle Stage 2

At Meazure Learning, we provide a host of services to our clients that encompass each of the Assessment Life Cycle stages. Below is a list of psychometric services that Meazure Learning offers specifically for **Stage 2: Item Development:**

Service	Description
Item development target planning	In order to have a valid, defensible examination, it is crucial to have a good understanding of how many items are required for each competency/competency area, based on a thorough and well-vetted examination blueprint. At Meazure Learning, we'll consult 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.). We can work with you to set item development targets and develop sound item development project management strategies.
Item development workshops	Item development is a careful and multifaceted process. Writing high-quality examination questions that conform to psychometric best practices requires specialized training and expertise. At Meazure Learning, we offer both face-to-face and remote training workshops in item development. We'll work hard to ensure that your item writers have the necessary knowledge and skills to produce the highest-quality items to fill your item bank.

Service	Description
Customized item development guides	It's not uncommon to have item writers continue to work on the development of new iterations of an examination for years; and sometimes, it's not feasible to provide ongoing item development training. To address these issues, at Meazure Learning we offer customized item development guides for our clients. These provide best practice general guidelines regarding item development; as well as specific guidelines that are tailored to the client's domain area.
Professional item editing and review	After raw items are produced, they need to be carefully vetted by individuals who are skilled in professional item development. At Meazure Learning, we offer professional item editing and item review services. Items are carefully reviewed to ensure content domain appropriateness, conformity to psychometric best practices, and correct grammar, punctuation, structure, and formatting. We'll work hard to make sure that the items you put forward your test-takers are all of the highest quality.
Professional item development	Finding subject matter experts who can spare the time required to create large numbers of items can be a challenging task. To help you meet your item banking needs, at Meazure Learning we offer professional item development services. This includes item cloning and automated Item generation. By using these services, we can help bolster your item bank without sacrificing the defensibility of the item development and review process.
Translation services	Many of our clients require their examination questions to be translated into multiple languages. At Meazure Learning, we offer professional translation services, which take items in a source language and translate them successfully into one or more target languages as required.