



Assessment Development: The Circle of Life

by Nikki Eatchel, Chief Assessment Officer, Scantron Corporation

The assessment world is always evolving—new theories, ideas, and approaches are regularly introduced. How can educators create and maintain an assessment environment that truly measures student achievement? The key is having a stable strategy that follows a continuous improvement cycle.

What does a typical cycle look like?

Assessment development is a continuous cycle with a defined starting place, but not a defined ending. Although your assessment development cycle may vary in complexity, typical cycles include some or all of the following phases.

Program Design

The program design is the primary strategic phase. This is where you examine your goals and determine how your overall assessment program can help you meet those goals. This planning phase involves determining how you are going to assess the whole student with testing, observation, performance-based tasks, student-created artifacts, and other educational tools. You need a combination of all of these components to accurately measure students across all of the hierarchical levels of a cognitive domain.

For example, do you want to include computer-adaptive testing? Adaptive testing tailors the test experience to each student and provides content that best meets their individual ability level. This testing model vastly reduces the “frustration zone” (where low-performing students see items too far above their current level of understanding and become frustrated with the test) and the “boredom zone” (where high-performing students already know the material and become bored with the test). In either case, you may get results that are difficult to interpret—designing an assessment program that mitigates these common testing issues can make your educational and assessment strategy stronger. Further, you will likely want to plan to measure skills attainment over time (longitudinal analysis) in addition to a snapshot of current grade-level proficiency.

A good approach can be to start with the analytics: What questions do you want or need to answer and what data will you need to answer those questions? Will you need to aggregate or disaggregate the results by demographics or other student groups? Do you want to compare student performance against participation in certain programs or identify outside factors (attendance, primary language spoken in the home, participation in extracurricular activities, etc.)? By identifying critical analytic factors at the beginning, you can design an assessment program that provides the answers you need to target instructional activities and improve student achievement.

Assessment Blueprint

The assessment blueprint (sometimes called a test development plan) lays out the exam and item-level details needed to meet the overall goals defined in your program design, often broken out by grade levels, subjects, and other assessment specifications.

Because assessment and education approaches constantly evolve, new standards and assessment models bring new challenges. Assessment teams must understand the implications of the ever-changing environment and the corresponding adjustments needed in their development processes—and sometimes even in the very way they think about assessments.

For example, the increasing influence of Norman Webb's Depth of Knowledge (DOK) and its reflection into the Common Core State Standards presents an interesting challenge around assessing students to determine their performance at varying cognitive complexity levels. Common Core standards often articulate necessary skills at a DOK Level 3: Strategic Thinking. Many districts and assessment teams naturally react to these requirements by focusing on the development of DOK 3-style items, with the goal of increasing assessment rigor.

However, if assessments include only items that measure a DOK 3 level of comprehension (even accounting for being grade-level appropriate), that exam design will limit educators' ability to drill down to a deeper diagnosis of how to help students who have not yet mastered the skill at that higher level of cognitive complexity. Knowing the current level of understanding for struggling students drives how you adjust instruction to move them forward. Meeting classroom goals such as this is a big part of a successful program design and assessment blueprint development.

Further, how much do you want to incorporate technology-enhanced items (TEIs)? While these items are certainly exciting to plan and use, they can require more development time—and tests that use them *must* be delivered online, removing your option to support both paper and online versions of the test. Not all districts, and not all schools within a district, are at the same level of technology readiness. Nor are all students ready to interact with technology in a way that truly demonstrates their mastery of the subject. Everyone involved in the assessment process has the challenge of balancing the needs of the available facilities and the students against the desire to use the "latest, greatest" item types.

These are just some examples of the many complexities inherent in developing a successful assessment program. A successful assessment team needs to understand these nuances and be prepared to create strategies to address them.

Item Development

The item development process identifies and creates the specific questions that will appear on assessments. This phase also includes researching reading passages and creating any images or other media required by the test items.

Whether you're developing informal quizzes and short assessments or more formal benchmark or interim assessments, the development process needs to produce fair and valid items. Successful items should support your assessment blueprint, track to grade-level standards, deliver multiple levels of difficulty, be free of bias, and accurately identify student achievement. In addition, passages and illustrations within items must be appropriate to the grade level or skill and properly licensed for use. Consider item option rationales during this process, as well, to allow for the development and documentation of supporting statements for both the correct item answers as well as the incorrect option choices.

To return to our previous example, a particular blueprint might specify a requirement to include certain percentages of items at specific DOK levels. That requires item developers to not only know their subjects, but also be able to deconstruct specific skills to determine how to measure those skills at all DOK levels. Remember, the goal of item writing is to develop an item that is the most direct measure of the intended learning outcome with minimal ambiguity. Every item should provide the best opportunity for students to display their knowledge.

If you plan to incorporate TEIs, it is important to evaluate your available staff to determine if you have all of the available resources to develop the different elements of the item (such as incorporated videos or animations). If your team members do not have the full set of skills and experience necessary to develop the items (or if they do, but do not have the available time needed to dedicate to the project), there are options available for purchasing or licensing some or all of the content you need. In addition to people resources, you must also ensure that you have hardware and software resources available for the development process as well (e.g., computers/laptops, content management tool, picture and video editing software, etc.).

As with all steps in the development process covered within this article, these are just some examples of the knowledge and abilities needed in item writers and the resources needed to fully develop and manage an item bank. The item-writing team must also attend to best practices regarding stems and answer choices, effective passage development and identification (including determining whether passages or multimedia items will require copyright permissions), appropriate graphic creation, and a host of other complex activities.

Item Review

Reviewing items ensures they are consistent, contain only necessary and appropriate content, support the right grade or difficulty level(s), and are free from bias. In addition, some teams determine scoring points or rubrics during this phase.

Item review sounds simple, but it is far from it. A strong review process can make or break the success of any assessment. Many assessment teams use a group of specialists whose sole duty is to review newly created items for a variety of factors, including but not limited to the following:

- Grade-level appropriateness of question stems, answer choices, passages (including readability score analysis)
- DOK or Bloom’s or other taxonomy appropriateness
- Gender, cultural, or socio-economic bias, including:
 - Images
 - Character names in passages, question stems, and answer choices
 - Question stems or answer choices that include or depend on a stereotype or are otherwise offensive
 - Topics and/or terms that are less likely to be familiar to particular gender or ethnic groups
- Proper grammar, punctuation, etc.
- Consistency in writing style
- Ambiguity (where it is not intended—some question stems or answer choices may need to be ambiguous because of what they are measuring)
- Appropriate documentation of rationales per item option

Item review teams require staff with sufficient time to thoroughly review the items, and ample experience with the factors listed above.

Exam Form and Bank Creation

The process of assessment development assembles completed items into the test forms or assessment banks you intend to present to students. This phase may include field testing of some or all items. It may also include operational testing to ensure your assessment delivery mechanisms run smoothly and efficiently.

Assessment developers must ensure that tests are neither too long nor too short, that they contain the necessary balance of item types and DOK levels, that the appropriate standards are covered, that various item and test versions are statistically and content equivalent, and much more.

Further, if assessment developers are responsible for field testing, they must select the item sets to be tested, identify a representative student population, schedule the test, review and analyze the results, and make changes to the assessment accordingly. If you choose to perform operational testing to ensure your systems can support your testing plan, the team must plan many of the same actions as they do for a field test, plus understand how to create tactics to address the inevitable operational challenges that arise.

Scoring

During your program design and item development phase, decisions impacting exam and item scoring need to be determined. A few considerations include:

- Will you be delivering fixed forms or using a computer-adaptive approach?
- What are the assumptions around scaled scoring?
- Will there be open-ended questions on the assessment?
- How will open-ended questions be scored? Who is responsible for rubric development?
- Who is responsible for scoring of constructed-response items? Will that be an external vendor or educators within the district/school?
- Will multiple-choice items be dichotomous, with only one correct answer?
- Will TEIs be included on the assessment?
- Will all items be weighted the same, or will some items be worth more than others?

Speaking of item weighting, do you plan to weight *assessments*—for example, an end-of-term test is worth more to the student’s grade than a unit, or a Spring end-of-term test is worth more than a Fall test? If you plan to use some of the same items on multiple tests, say, because your Spring test is inclusive of Fall items, you may want to consider weighting the test.

If you are thinking about TEIs, you must also consider how these items will be scored. One of the major unanswered questions about TEIs is whether they truly provide improved opportunity for students to demonstrate their knowledge in a statistically valid and reliable way. Or, asked differently, does a TEI item actually increase the accuracy of student evaluation over and above what is measured in a traditional selected-response item? More research needs to be completed to evaluate a variety of challenges around TEIs, such as reliability and interpretation, including how to determine whether a student missed the question because they don’t know the answer or because they don’t know how to use the technology. Including TEIs for technology’s sake alone can impact your results in unexpected ways.

Analysis and Reporting

The ultimate purpose of assessment is to gather data and information that you can use to drive instruction. Psychometric analysis and reporting is key to turning your assessment data into meaningful information.

A well thought-out analysis and reporting plan can deliver the best, most actionable results, completed as part of your assessment delivery activities. For example, understanding the effectiveness of items and tests in your assessment environment can help to demonstrate the value of your assessments in measuring student achievement. This is often the key to acquiring increased funding and other opportunities. Many districts struggle to find the time and expertise to conduct these analyses in addition to their other critical duties.

This phase may also include special research projects to determine the best use and interpretation of the data, including whether your assessments can be used to predict student performance, the impact that instruction and assessment have on student outcomes, or the potential instructional interventions that may be necessary in the future.

Maintenance

Great! You've done your due diligence, developed your program, your blueprint, your items, and your assessments. You've determined delivery methods, item and test scoring and weighting, and you're ready to deliver your first test (plus you'd probably like to take a nice nap because you're exhausted). You can dust off your hands and sit back, waiting for results to come in, knowing you won't ever have to do this again, right?

Sadly, no. To ensure academic rigor and reliable results over time, you must maintain the assessments as well. This is called the assessment development **cycle** for a reason! Doing your due diligence for your assessment program is a little like painting the Golden Gate Bridge: Once you finish, it's time to go back to the other side and start all over.

The only constant in education is change: approaches, theories, requirements, and other factors that affect your assessments regularly evolve and grow. This means you need to plan to address these changes in your assessment program. You may need to recruit new subject-matter experts (SMEs) to craft new content, review existing content, or ensure item banks are aligned to new and evolving standards. You'll want to pay attention to trends and upcoming required changes and forecast your development needs so you're not caught by surprise.

In addition, just like everything else in life, items can be used past their peak if not continually monitored. When used over an extended period of time, items can potentially fall outside of optimal exposure levels. Items that have those larger exposure levels can sometimes experience item drift (become easier) and can potentially contain content that is no longer as relevant as when it was first developed. You should plan to refresh your content regularly—which also means you can use topical items that refer to current events or social trends more readily, thus engaging your students in the assessments more thoroughly and relevantly. A key part of the refresh process involves overall item bank review and assessment development forecasting to determine the areas in which new development is needed to support your current and future assessment needs.

Performing maintenance tasks and analyzing your assessments for continual improvement is a critical step in ensuring you are getting the most accurate results to measure student achievement.

How can you implement this cycle to suit your needs?

The process of developing and maintaining quality assessments is highly individual to each district, as is deciding who should be on the team and what expertise they have—or need.

Some districts follow a process that includes every step within the cycle described earlier. Others choose to focus on certain steps, or even only on certain activities within each step. Often, these decisions come down to resources: time, available staff, and existing expertise. Within these decisions lives a spectrum of options.



In-house development

In this model, districts decide to completely support their assessment program with only in-house resources. Teams are assembled to plan, develop, review, and maintain assessment items and test forms.

Pros

- In-house team has complete control of process and timing.
- Team is already familiar with district culture and assessment needs.
- Teacher buy-in is more likely.

Cons

- May not have staff members already in place with requisite expertise.
- Team members may not be dedicated solely to assessment project efforts, which can lengthen project timeline.
- Maintenance activities are difficult to schedule and complete.

Supplemental assessment support

In this model, districts choose to use a combination of efforts by internal teams and outside partners. Some elements are completed internally, while others are outsourced to one or more partners.

Pros

- District can maintain control while leveraging internal and partner resources appropriately.
- District team often increases capacity and expertise by working closely with partner.
- District team is already familiar with district culture and assessment needs.
- Partner has readily available team with requisite expertise

Cons

- District may not have staff members already in place with requisite skills and knowledge of internal elements.
- In-house team members may not be dedicated solely to team efforts, which can lengthen project timeline.
- Maintenance activities are difficult to schedule and complete.
- Communication between partner and district teams must be carefully managed.

Full-service support

In this model, districts decide to work with a partner (or partners) to complete the majority of activities within the assessment cycle. The district provides guidance and the partner creates the blueprint, develops and reviews items, and completes all of the tasks necessary to deliver high-quality assessments.

Pros

- No significant burden on existing resources.
- Partner has readily available team members with requisite expertise.
- More efficient process using partner with deep experience in assessment development.

Cons

- Some loss of direct control.
- District leaders must be available to advise about district culture and assessment needs.
- Without teacher involvement, buy-in may be more difficult.

Many districts want or need to make assessment content their own, for a variety of reasons, including ensuring teacher buy-in. Once you know what implementation approach will suit your district best, the next step is to look for a partner who can design support to meet your exact desire for involvement, ensuring you remain in the driver's seat.

What kinds of help should you expect from a partner?

Whatever cycle segments you adopt, and whichever model you choose, look for a partner who can support your efforts. In-house development teams can benefit from professional development workshops or expert consulting hours. Full-service support partners should always work closely with your district to ensure your assessments meet your needs today and anticipate and fulfill your needs for tomorrow.

Supplemental support partners should be able to step in at every stage and provide the specific assistance needed. Since the only constant is change, and the model you decide to follow for one round of the cycle may not work the next time, consider looking for a single partner who can help you today *and* tomorrow.

See the checklist at the end of this article for sample questions to ask when choosing a partner.

How can Scantron help?

Scantron has an extensive track record of providing assessment services to help thousands of customers succeed in their assessment programs. We were one of the first partners to align our existing item banks to the Common Core State Standards—and we were one of the first to provide item and assessment banks designed *from the ground up* to fully support those standards. We've helped districts with workshops and consulting on a variety of assessment topics and content development processes, as well as working with customers to provide high-quality, aligned content. We've validated educator assessment efforts and demonstrated positive impact to student

learning with research studies ranging from determining growth targets to predictive validity studies.

One critical challenge is to develop an assessment plan that fulfills all of the needs of a district's educators. This process might include developing a plan for custom assessments that includes creating an assessment blueprint and effective test items to accurately measure student achievement. It might also include aligning district-created items to a particular set of national or state standards. Scantron can develop these deliverables using the same proven assessment development process we follow to create our own item banks and assessments—or you can choose to incorporate existing Scantron materials into your plan. If you want to play a more active role in assessment creation, Scantron can assist with a variety of professional development activities that will instruct educators in how to create test blueprints and develop assessment items. Last, if linking district test results to state test results or performing other studies is important, Scantron's psychometric and research team can help.



Another key element is providing support to build district team capacity and inform Professional Learning Communities (PLCs). Scantron can develop your assessment elements—and we can provide consulting and workshops on a variety of topics, including:

- Blueprint development
- Item development plan
- Item writing and review
- Bias and sensitivity review
- Test form creation and delivery
- Statistical analysis and interpretation of results
- Custom research projects

Whatever the assessment assistance you need, Scantron has the products, tools, and expertise to help you ensure that you have the right program for your students. Our comprehensive suite of assessment services helps you make the most out of your assessment development process. Whether you want to do it yourself with in-house experts and just need a little support or whether you want a partner to guide you through the process, Scantron can meet you where you are and help get you to where you want to be.

Contact Us:

1.800.722.6876

www.scantron.com

Scantron Corporation, 1313 Lone Oak Road, Eagan, MN 55121

Don't Just take our word for It! Some quotes from our Clients:

"We began by having Scantron use our blueprints to build test items. We have learned so much from [Scantron's] expertise and guidance. The research provided by their psychometricians has been key to our success!"

Mary Lou Caldarera,
Supervisor of Assessment, Calcasieu Parish, LA

"[Scantron] not only helped write, edit, and review test items for our core content areas, but they've helped orchestrate the field testing and the psychometric review of these items so that they pass a single parameter model.

Thea Andrade,
Assistant Superintendent of Curriculum, Assessment, and Instruction,
Phoenix Union High School District

"Scantron has provided the tools needed to validate student learning and support teacher decisions to move students to the next Performance Level when they are ready."

Dr. Oliver Grenham
Chief Education Officer, Adams County School District 50



Questions to Ask Your Assessment Partner

Process	<ul style="list-style-type: none"> • What are the steps within your assessment development process? Are they supported by industry standards? • What deliverables are included? Which cost extra? • Which individual elements of our assessment development cycle can you support?
Experience	<ul style="list-style-type: none"> • We're following [insert your foundational approach, e.g. PARCC, SBAC, custom Common Core, custom, state, etc.]. What do you know about that? How does that change your development approach? • With whom have you worked before?
Content	<ul style="list-style-type: none"> • Do you already have items or assessments that might fit our needs? • What level of subject-matter and assessment expertise do your team members have? • Have the items been field tested? • What types of reviews have these items gone through? Have they been statistically analyzed to assess performance? • What alignment capability or reporting do you offer? • What kind of predictability or other results are your customers seeing when they use the assessments including these items?
Building Capacity	<ul style="list-style-type: none"> • What consulting or professional development services do you offer regarding creation of quality assessments and analysis of data so that we can develop our own team's expertise? Can you offer that support for our Professional Learning Communities (PLCs)?
Psychometric Services	<ul style="list-style-type: none"> • What types of analysis or research studies do you recommend? Do you have sample reports you can provide? • For whom have you done research studies? How did the results benefit the district?
Ongoing Assistance	<ul style="list-style-type: none"> • What types of support can you offer for regular assessment checkups—annual content review, replenishment planning, alignment updates, etc.? • What about professional development refreshers for internal teams and PLCs?
Software Independence	<ul style="list-style-type: none"> • Can your vendor offer you assessment services independent of software or product purchases? Or do services have to be combined with product?
Return on Investment	<ul style="list-style-type: none"> • How will your assessment partner help you calculate your current ROI and projected ROI with the available assessment models (in house, supplemental, full service) or the ROI at any of the life cycle stages?



**INFORM INSTRUCTION
TO IMPACT STUDENT
ACHIEVEMENT TODAY!**

For a free consultation to meet your academic goals, call **800.722.6876** or visit us at www.scantron.com/k12.