



Digital Test Delivery:

Empowering Accessible Test Design
to Increase Test Validity for All Students

MICHAEL RUSSELL





Executive Summary

The development of digital test content, coupled with computer-based test delivery, provides an important opportunity to improve the accessibility of test items. By applying principles of accessible test design, the next-generation assessment systems will deliver more valid inferences about student learning based on test scores for all students. Rather than developing assessment content for the general population of students and then making post hoc changes to accommodate the needs of subgroups of students, accessible test design provides a framework for making careful decisions about the methods used to tailor test administration to maximize the measurement of targeted constructs for each student. In turn, the Accessible Portable Item Profile (APIP) standards provide a tool for implementing accessible test design. The APIP standards empower next-generation assessments to solve three challenges. First, APIP provides a structure for specifying and storing the access needs of each student. Second, APIP provides a structure for augmenting item content with a variety of supplemental and alternate accessibility information designed to ensure that a test item functions properly for students with a variety of access needs. Third, APIP provides specifications for developing test delivery systems that can use a student access profile to tailor the provision of access tools (such as magnification, color contrast, masking) and the presentation of supplemental accessibility information (audio, Braille, tactile, or signed versions of item content). Collectively, the tools provided by APIP enable next-generation assessments to capitalize on the flexibility of digital technologies to maximize test validity for all students.

Russell, Ph.D. is Vice President of Innovation at the Nimble Innovation Laboratory, a division of Measured Progress that specializes in computer-based assessment solutions that are accessible for all students including those with disabilities and special needs. Dr. Russell is also an Associate Professor at the Lynch School of Education, Boston College.

This paper was produced in partnership with Arabella Advisors. Arabella Advisors is a philanthropy consulting firm supporting the efforts of individual, family, institutional, and corporate donors worldwide. We are committed to unbiased analysis that helps donors support issues and nonprofits with confidence. Our expertise and insights transform philanthropic goals into results. For more information on our firm, please visit our website at www.arabellaadvisors.com.



RECOMMENDATIONS

1. Adopt the Accessible Portable Item Profile (APIP) standards or a similar set of item profile standards that provide a comprehensive accessibility solution
2. Establish a best practices working group to develop business rules for implementing the APIP (or others as designed) accessibility elements
3. Require that an access needs profile be formed for each student
4. Carefully define the construct intended to be measured by each item or task
5. Require that accessibility information be embedded in each test item
6. Require that the test delivery system is compliant with the APIP standard or a similar set of item profile standards that provide a comprehensive accessibility solution



It is only during the past decade that students in the margins have been fully included in large-scale educational testing programs.

Introduction

For the past 40 years, the educational testing community has struggled to accurately and validly measure the achievement of students in the margins, including English language learners and students with disabilities or special needs. For a long time, many of these students were simply excluded from testing programs. Over time, however, legal action and advocacy led to the provision of test accommodations that allowed students in the margins to participate in testing programs. Initially, test accommodations were highly controversial and many programs either excluded or flagged scores for students who were provided test accommodations. It is only during the past decade, with the adoption of the No Child Left Behind Act, that students in the margins have been fully included in large-scale educational testing programs. Nonetheless, the equity and quality with which test accommodations are provided, the expense associated with the provision of test accommodation, and the effects that test accommodations have on the validity of test-based inferences continue to raise concerns.

The adoption of computer-based testing holds promise for addressing some of these concerns. In fact, the flexibility with which the delivery of digital content can be tailored for each individual student provides a unique and powerful opportunity to increase test validity for all students. This paper explores how digital test delivery empowers test developers to purposefully design items and tasks that maximally access the intended construct within each student. It begins by reviewing how test items are designed to function and then explores how digital content and digital delivery can be designed to create flexible test taking experiences that maximize the accuracy with which intended constructs are measured.

The Challenge: Designing Tests that Work for All Students

A test is designed to measure a specific construct or set of constructs. For ease of reference, the specific construct or set of constructs measured by a test is referred to as the intended construct. To provide a measure of the intended construct, a test is composed of a set of items or tasks or both. Each item or task is specifically constructed to create a measurement experience that requires the examinee (for purposes of this paper, we will assume the examinees are students) to apply the intended construct. Each measurement experience involves a carefully crafted three-step process. During the first step, the student is presented with information (or content) that is designed to establish a problem or question that stimulates the construct of interest. During the second step, the student is provided an opportunity to interact with content contained in the item or task while applying the intended construct. Since the application of the construct of interest cannot be directly observed, the third step requires the student to produce a response that represents the product or outcome of the application of the intended construct. It is through this three-step process that a test item or task attempts to access the intended construct as it currently operates within the student. Through observations of the student's application of the intended construct provided by multiple items or tasks, a test allows the outcome of each observation to be accumulated to produce a test score that is used to make an inference about the extent to which the intended construct operates within the student.

When creating a test item or task, item writers generally design items that function optimally for the general population of students. When considering the general population of students, several assumptions about students are often made. As an example, it is often assumed that the student does not experience any chal-

lenges with fine or gross motor skills, does not have any visual or auditory needs, is able to read near or above grade level, and is proficient in English. It is also assumed that executive functioning skills are adequate to allow the student to remain on task, identify key words or phrases in an item, work fluidly with multiple pieces of information, and break complex tasks into discrete elements. Finally, it is often assumed that the student is able to produce responses using either a pencil, keyboard, or mouse. While these assumptions may apply for the majority of students, they tend not to hold for students in the margins.

For these students, the knowledge and skills that item writers typically assume test takers possess present barriers to accurately and reliably accessing and measuring the intended construct. As depicted in figure 1, each phase of an item's functioning presents potential access barriers. As an example, most test items present content that is designed to stimulate the intended construct in narrative English form. For students who are English language learners, have visual impairments, have difficulty decoding text, read below grade level, or are unfamiliar with words or phrases that appear within an item (e.g., names of people or objects), the text-based narrative format of an item stem may not stimulate the intended construct as designed. Similarly, some items and tasks that present multiple pieces of information, contain multiple steps, or mix the format in which information is presented (through narrative text, tables, graphs, graphics, or video, among others) may be difficult to interact with for students who experience challenges with executive functioning, task prioritization, or information processing. Finally, items that require either written responses or the use of a mouse to select an option or manipulate digital objects may yield inaccurate responses for stu-

dents with physical disabilities or limited fine or gross motor skills. Across these three phases of an item's functioning, inaccurate stimulation of the intended construct, difficulty interact-

ing with item content, or difficulty producing responses that accurately reflect the application of the construct produce challenges to an item's ability to measure the intended construct.

Test Accommodations: Why Post Hoc Changes Are Inadequate

For paper-based tests, accessibility challenges have traditionally been addressed by providing test accommodations. By definition, test accommodations are changes to test administration that address a specific access need. Traditionally, test accommodations are considered after a test form is developed and involve creating different versions of the test form (e.g., a large-print or Braille version), allowing an access assistant to translate a test into another form (e.g., reading an item aloud or presenting item content in American Sign Language), or making other changes to test administration conditions (e.g., allowing a student to use masks or reading guides, allowing a proctor to assist with managing test materials, or allowing the student to record answers in the test booklet). As noted above, this approach to enhancing access through post hoc test accommodations presents at least four challenges.

1. Creating alternate versions of test forms increases the expense for a testing program.
2. Because post hoc alternate versions are often made in the absence of item writers, the construct measured by an item may be changed inadvertently.
3. In addition to adding cost to a program, relying on access assistants to make post hoc translations of item content leads to differences in translations and the quality with which translations are delivered to students.
4. Inequity in the resources (both physical and human) available across schools produces inequities with respect to the opportunity to alter test administration conditions to meet specific needs.

As a result of these challenges, opportunities to receive test accommodations and the quality with which accommodations are provided vary across schools. Collectively, these challenges limit the effectiveness of traditional post hoc accommodations for improving the ability of test items to access the intended construct.

Digital test content and delivery provide an important opportunity to overcome these challenges. This opportunity results from two features of digital content and delivery: the ability to embed additional accessibility information into digital content files as an item is developed and the ability of a digital delivery system to selectively present subsets of that information to individual users based on their specific need.

Default and Supplemental Item Content

As Mislevy and his colleagues explain, different representational forms can be used to present item content to a student.¹ To enable a student to recognize and process content, the form used to present that content may need to be tailored based on the student's representational form need. As an example, content presented in print-based form will not adequately stimulate the intended construct for a student who is blind. However, when that same content is presented in Braille, the content is able to access the intended construct within a student who is a Braille reader. Similarly, for a student who is deaf or hard of hearing, content presented in audio form may not adequately stimulate the intended construct. However, when presented in a signed form, the content is able to stimulate the intended construct within a student who communicates in sign. Forms of alternate representation include reading content aloud, presenting text-based content in sign language, Braille, tactile representations of graphical images, symbolic representations of text-based information, narrative representations of chemical compounds (e.g., "sodium chloride" instead of "NaCl") or mathematical formulas, and translating to a different language.

Related to the notion of alternate representations are distinctions among default content, alternate content, and supplemental content. Default content is item information that is presented to a student who does not have defined access needs. Typically, default content includes text, graphics, and/or tables that form the item as developed for the general population of students.

Alternate content presents a different version of the item to students with specific needs. In essence, some or all of the original content is replaced by other content. Examples include

using a translated version of an item (e.g., Spanish instead of English) or wording that is more easily understood (i.e., simplified English).

In contrast, supplemental content provides additional content to address access needs. For example, text might be supplemented with an audio, Braille, or signed form of the content. Similarly, to assist in identifying important aspects of content, supplementary information may be presented, such as highlighting of key words, translation or definitions for key words, or flags that point the student to key information.

As is described in greater detail below, digital content files allow item developers to specify default content and supplemental content for each item or task. In addition, a digital content file can provide a pointer to separate files that contain alternate versions of the default content (e.g., Spanish translation). Within a digital file containing an alternate version of an item, default and supplemental content for that alternate version can be provided (e.g., default content presented in Spanish accompanied by supplemental content specifying how to read aloud that content in Spanish). It's important to note that specifying default, supplemental, and alternate item information during the item development phase allows item writers to carefully consider whether supplemental and alternate information alters the intended construct measured by the item. If it does, the item writer can either modify the supplemental or alternate information or determine that supplemental or alternate information cannot be provided. By making these decisions during item writing, a test program can ensure that the same high-quality supplemental and alternate information is available for all students while assuring that the item still measures the intended construct.

Digital Item Delivery

Capitalizing on the flexibility of computer-based technologies, computer-based test delivery interfaces can tailor the presentation of default, alternate, and supplemental item content, interactions with that content, and response modes based on each individual's needs. To do so, developers should employ principles of universal design when creating systems that can personalize the testing experience based on each individual student's needs.

The concept of universal design focuses on “the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design.”² Rather than creating a single solution, universal design has come to embrace the concept of allowing users to select from multiple alternatives. As Rose and Meyer emphasize, “Universal design does not imply ‘one sizes fits all’ but rather acknowledges the need for alternatives to suit many different people’s needs ... the essence of [universal design] is flexibility and the inclusion of alternatives to adapt to the myriad variations in learner needs, styles, and preferences.”³

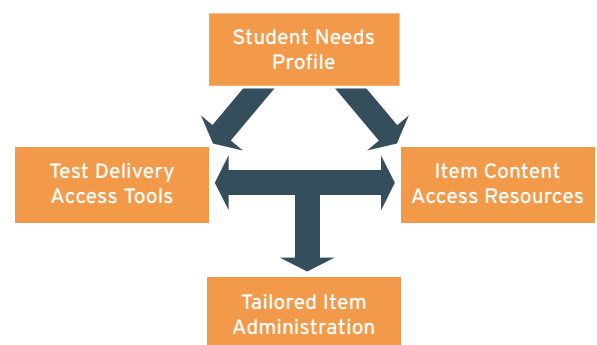
Technology allows developers to apply principles of universal design to educational assessments in a way that improves access for all users. When building a universally designed educational assessment, three important aspects must be considered:

1. To access the intended construct within each student, it must be acknowledged that a single default version of an item is not adequate. Instead, supplemental and alternate information that is carefully designed to access the intended construct within specific subgroups of students must be embedded into item content.
2. A user access needs profile must be developed for each student, and that profile must

specify the method(s) of presenting, interacting, and responding to item content that are expected to best access the intended construct. The profile defines access needs for a given student and indicates which tools or representational forms should be made available for that student. The profile might also give specific settings, such as magnification levels, color contrasts, or default representational forms preferred by the student. Once defined, the access profile interacts with both the delivery interface and the item content.

3. The interface used to deliver items and tasks must be able to interact with each student's access needs profile and with the default, supplemental, and alternate information specified in each item. The interaction with the delivery interface focuses on specific tools or features embedded in the interface, activates whichever tools and features are defined in the profile, and, in some cases, controls the exact settings for those tools and features. The interaction with the item content focuses on which of the specific representational forms embedded in the item should be presented or activated for a given student in order to meet that student's specific need.

FIGURE 1:
ACCESSIBLE TEST IMPLEMENTATION MODEL



As depicted in figure 1, these three components must be designed to work together to

tailor test delivery to meet the specific access needs of each student.

Accessible Test Design

Accessible test design addresses item content, representational forms, test delivery interface, and access profiles by specifying methods for flexibly tailoring an item so that the influence of non-targeted constructs is reduced for each individual student. Depending on a student's access needs, flexible tailoring may require an adaptation to the presentation of item content, the interaction with that content, the response mode, or the representational form in which content is communicated.

Adapted presentation may require item content to be presented in manner that assists the intake of information, such as magnifying or adjusting the contrast level with which item content is presented. Adapted interaction may require changes to the conditions under which a student applies the targeted construct, such as decreasing distractions by masking content, providing auditory calming, or highlighting key content within an item. Adapted response may require the student to use different types of tools to produce responses, such as a speech-to-text or an assistive communication device. Tailored representations may require item content to be presented using a different representational form, such as Braille, sign, audio, an alternate

language (e.g., Spanish), or using simplified vocabulary.

Providing these adaptations and tailored representations in a consistent manner requires careful thought during the item development stage. To assure that adaptations and tailored representations do not negatively influence the validity of inferences based on the resulting test score, item developers must specify supplementary and alternate information associated with each item and assure that these alternate representations do not alter the measure of the intended construct. Providing adaptations and tailored representations in a consistent and valid manner also requires assessment programs to take a systems approach to accessibility. When designing a test delivery interface, an assessment program must specify the variety of accessibility tools and features that may be required for specific students. Finally, the test delivery system must be able to integrate student access information, item accessibility information, and interface accessibility tools and features to tailor item delivery to meet access needs for each individual student. Accessible test design requires a comprehensive model for test design and administration.

Accessible Test Design: An Example

For the past two years, the New England Common Assessment Program (NECAP) has implemented principles of accessible test design for its operational science tests and for its grade 10 operational mathematics and English language arts test. As part of this early adoption effort, the NECAP⁴ states have allowed students for whom schools felt default item content did not accurately access the intended constructs to use a universally designed test delivery interface to perform the test. For these tests, digital versions of test items were developed. For each item, default and supplemental (but not alternate) item content was specified. Specifically, the supplemental item content focused on providing audio access to print-based content for students who had decoding challenges and for students who had vision needs. In addition, a universally designed test delivery interface was used to tailor the presentation of, interaction with, and response to items. Specifically, the test delivery interface allowed several types of tailored delivery that included:

- Presentation of content: (1) magnification; (2) reverse contrast; (3) color tint; and (4) alternate foreground and background color
- Supplemental content: (1) audio presentation of narrative content; (2) audio presentation of graphics; (c) audio presentation for nonvisual users; and (d) tactile presentation of graphics
- Interactions: (1) answer masking; (2) custom masking; (3) breaks; and (4) auditory calming
- Response: (1) mouse; (2) keyboard; (3) tab-enter controlled alternate communication devices; (4) touch screen; and (5) Intellikeys


While this early effort provides a powerful example that accessible test design can be implemented at scale for operational tests, it has also identified several challenges to the adoption of accessible test design.

Challenges to Accessible Test Design

Perhaps the most important challenge to address focuses on developing a clear definition of the intended construct measured by an item. Without a clear definition, it is difficult to determine whether supplemental and alternate item content alters the construct measured by the item.

A related challenge focuses on the need for access experts to interact with item content experts throughout item development. It is unlikely that any one individual will have sufficient knowledge of the intended construct and the various types of supplemental and alternate

content required to allow the item to access that construct within all test takers. For that reason, item writing must involve an iterative team approach that includes a content expert and experts familiar with the specific needs addressed by the supplemental and alternate content (e.g., an expert in tactile representations for an item containing figures or graphs, or an expert in American Sign Language for an item that must be translated to ASL). While this may initially seem time- and labor-intensive, business rules for developing supplemental and alternate information can streamline this process.



Another challenge is that in order to adequately access the construct within each student, educators must make informed decisions about individual access needs and assign an appropriate access profile for each student. Since the concept of access needs and flexible test delivery is new to many educators, professional development and decision-making tools are required.

The development and use of a flexible test delivery engine that is able to interact with a student access profile and the default, supplemental, and alternate content associated with an item are also requisite. As new access needs are addressed by an assessment program and additional item information is incorporated into item files, it's important that the delivery engine be easy to adapt to incorporate these new techniques and tools.

Finally, there is a clear need for common expectations regarding the type of accessibility information that is included in an item file

and the behaviors that should result when a test delivery system interacts with that information. Again, because the concept of accessible test design and tailored item delivery is new, consensus has not yet been reached about how some access needs are best met (e.g., how auditory support for sighted students who require decoding assistance may differ from auditory support for students who are blind). An important step toward reaching consensus about these issues is to form working groups that focus on developing business rules for providing specific types of supports (e.g., tactile and Braille representations) and guidelines for creating scripts for auditory and signed representations (e.g., how to read or sign dates, exponents, chemical equations, and so on). In addition, because this approach of developing supplemental and alternate item content is new, a standard for tagging or coding that information is needed so that items are interoperable across delivery systems.

Accessible Portable Item Profile Standard

To meet many of these needs, the APIP Project has developed an open standard called the Accessible Portable Item Profile (APIP). Led by the Minnesota Department of Education, the APIP Project includes a consortium of eight states, IMS Global Learning Consortium (an interoperability standard-setting organization), and experts in testing, accessibility, and interoperability standards. Using the concept of accessible test design as a foundation, the APIP standard provides an open standard for specifying default, supplemental, and alternate item content, and for identifying the access needs for each individual student.

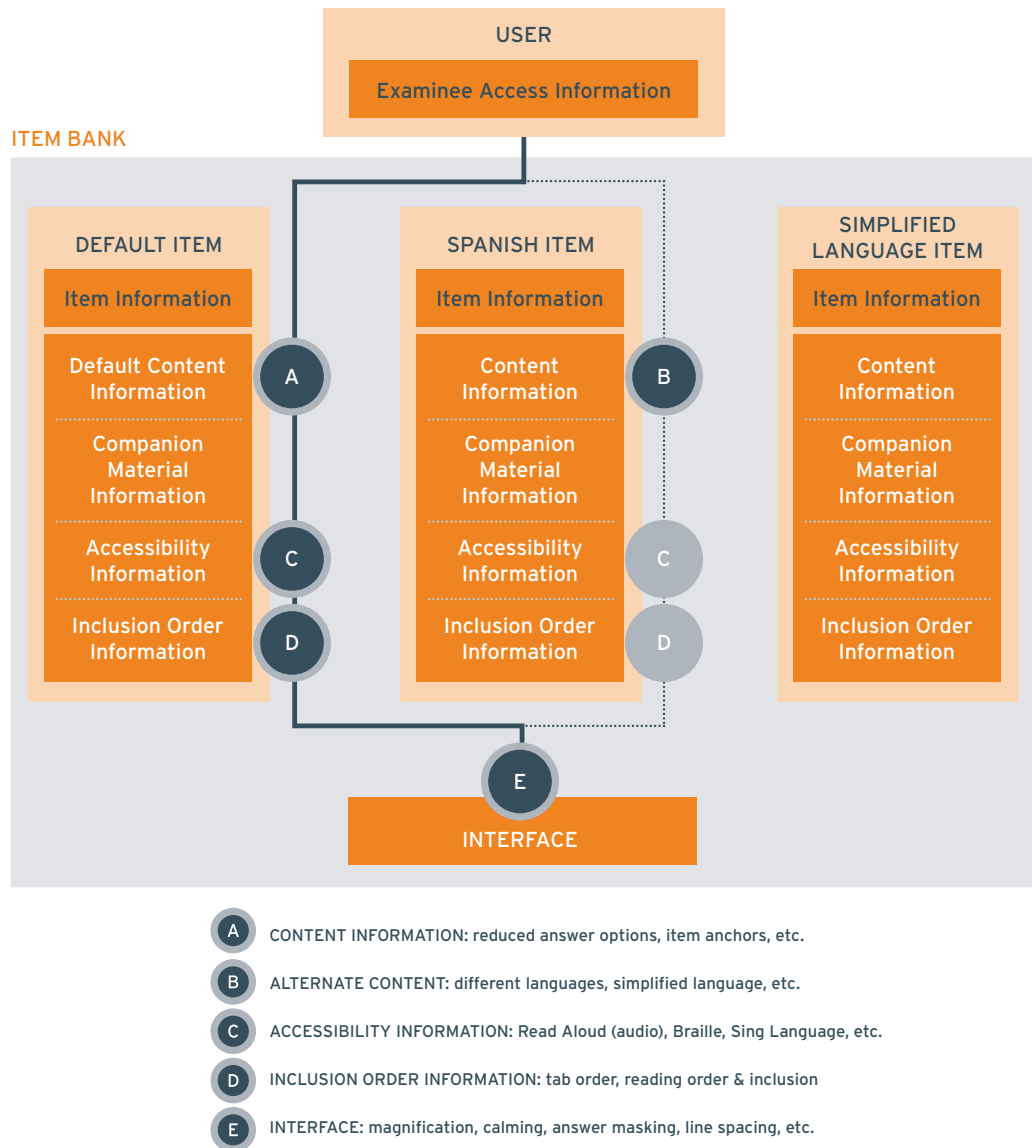
As depicted in figure 2, the APIP framework is composed of two information models that allow test delivery engines to tailor the presen-

tation of items to meet the access needs of each individual student. The first information model focuses on Examinee Access Information. During test delivery, the Examinee Access Information model performs two functions. First, it provides information that allows a test delivery engine to activate specific tools that tailor the presentation of item content to the student. These embedded access features may include magnification, alternate contrast, increased white space, and answer masking. Second, it provides information that specifies which accessibility information embedded within the item model is pertinent to the student. APIP allows item developers to place a variety of types of access information within an item, including specifications for how an item is to be pre-

sented in auditory, Braille, sign, or tactile forms. In addition, the item information model allows an item developer to point to alternate versions

of the item that are presented in an alternate language (e.g., Spanish) or in simplified English (e.g., with negatives removed).

FIGURE 2: APIP MODEL



As depicted in figure 2, the APIP item model has five components: Item Information, Default Content Information, Companion Material Information, Accessibility Information, and Inclusion Order Information. Each component is described briefly.

The Item Information component provides

meta-information about an item. Examples of meta-information include the domain, sub-domain, intended construct, intended grade and age level, item difficulty, item discrimination, and item exposure rate. In addition, the Item Information component contains information about alternate versions of the item for which

item content has been substituted to make the item accessible in a different language or as a simplified version of the item in the default language.

The second component, Default Content Information, provides information about the contents of the item that are to be presented to a student assuming no access needs have been defined for that student. This information includes the item type, the prompt, media associated with the item (e.g., figures, tables, and graphs), response options, correct response, and scoring rule.

The third component, Companion Material Information, provides information about materials and tools that the student is expected to work with while performing the item. These materials and tools may include such things as a reading passage, a primary document, a periodic table, ruler, protractor, and calculator.

The fourth component, Accessibility Information, provides information about alternate representations of default content. The types of alternate representations specified in this component include audio, signed, or tactile presentation of item content. The Accessibility Information component may also include specifications for scaffold supports or key word translations.

The fifth and final component, Inclusion Order, specifies the order in which accessibility elements are to be presented to a student with


a given category of access needs. The categories of access needs for which an inclusion order is specified include: (1) text-based audio; (2) graphic audio; (3) text-based and graphic audio; (4) nonvisual audio; (5) Braille; (6) American Sign Language; and (7) other sign language. Text-based audio access is specific to students who are able to view contents displayed on a screen but may need assistance accessing those contents. Often, students requiring text-based audio access read below grade level, have reading related disabilities, or the language in which item content is displayed is not their primary language. When audio forms are presented as core content, the item developer must specify the order in which content is presented. In addition, the item writer must consider whether all item elements are presented as core content in audio form. As an example, an item writer may opt not to present labels associated with a graphical element as part of the core content, or may opt not to read the contents of a table when the item is read from beginning to end. In these cases, item content that is not presented as core content may be accessed by the student on demand.

In sum, the APIP model empowers item developers and testing programs with a standard method for specifying the tailoring of items to meet specific accessibility needs and provides a foundation for accessible test design.

Looking to the Future

The Race to the Top (RTTT) Assessment Program holds promise to stimulate several advances in the field of assessment. Among those advances are the development of innovative item types and performance tasks that can be used at scale to assess complex, higher-order skills. Wide-scale adoption of adaptive testing is

likely. In the area of formative assessment, there is potential for diagnostic instruments that help educators identify what students know and don't know and that provide information about misconceptions or other underdeveloped reasoning that may interfere with a student's conceptual understanding.



Across all forms of assessment and types of tasks used to collect information about student knowledge and understanding, it will be important that instruments adequately access the intended construct within each individual student. Accomplishing this will require a proactive *a priori* approach to developing accessible assessment content. Given the large number of instruments and accompanying items and tasks that must be developed, as well as the new ways in which students will be expected to interact with item and task content, accessibility will be a major challenge. However, the development of such tools as APIP and flexible test delivery interfaces provide an opportunity to examine issues of accessibility during prototyping and early stages of development. By doing so, the field can be proactive in identifying potential issues that specific item types or item content may present for students with specific access needs. Early identification will then allow researchers and development teams to explore alternate item and task designs that may overcome these

access barriers. Or, in some cases, test designers and item developers may decide that the access need is so tightly intertwined with the intended construct (e.g., vision intertwined with using a microscope to locate and identify microscopic organisms) that the construct cannot be validly measured for students with that access need. Such decisions will have subsequent implications for the metadata associated with an item, the design of adaptive test engines, the generation of student scores, and ultimately the types of inferences that can be made about a student based on that score. While the flexibility that digital content and digital delivery offer with respect to tailoring item and task delivery to maximize access to the intended construct within every student may seem overwhelming, tools like APIP empower assessment programs to make important decisions throughout the test development process about what exactly is being measured, what is not intended to be measured, and how to tailor item and task deliver to maximize validity for all students.



Take a proactive *a priori* approach to developing accessible item and task content.

Recommendations

- Take a proactive *a priori* approach to developing accessible item and task content. Rather than repeating past practices that have raised questions about the extent to which test accommodations change the measured construct, create item writing teams that are knowledgeable about the intended constructs and about the provision of supplementary accessibility information that does not alter the construct measured by the item.
- Adopt the established Accessible Portable Item Profile (APIP) standards or a similar set of item profile standards that provide a comprehensive accessibility solution, and use it as a foundation for item development, student rostering, and test delivery systems. The APIP standards provide a structure for specifying and storing student access needs and defining supplemental and alternate item accessibility information. Test delivery systems, in turn, must be developed to read and interpret a student access needs profile and to tailor the provision of accessibility options based on that profile.
- Form a working group to develop best practices for implementing APIP standards, or a similar set of item profile standards that provide a comprehensive accessibility solution. The standards adopted should be powerful tools for specifying how content is presented to students with specific access needs. To assure that the tools are applied in a consistent manner that does not violate a measured construct, business rules must be developed to help guide decisions about how to present content to meet a specific need (such as how to read aloud exponents, dates, or scientific notation, or how to present simple and complex tables in auditory and Braille form, and so on). Ideally, the best practices working group will include representatives from both RTTT assessment consortia and will comprise experts in specific content areas and in specific access needs, as well as experts with experience in implementing accessible test design.
- Clearly define the construct intended to be developed by each item and identify access-related constructs that are not intended to be measured by each item. Without a clear definition of the intended

construct, sound decisions about the type of supplemental accessibility information embedded within an item are impossible. It is only with clear definitions of the intended construct that sound decisions can be made about whether supplemental accessibility information alters the measured construct or supports valid test-based inferences about student achievement of that construct.

- Take a team approach to item development that incorporates supplemental and alternate content into items. Provision of supplemental accessibility information requires an understanding of the intended construct and the varied accessibility needs of students. No one person possesses all of this knowledge. Hence, a team approach is needed to develop and evaluate the quality of accessibility information provided for each item.
- Be prepared to accept that some items may not be able to be made accessible for students with a specific access need and develop test specifications or adaptive algorithms that take this limitation into consideration. There will be intended constructs that overlap with access needs (e.g., using a microscope to identify microscopic objects overlap with vision). When these cases occur, it may not be possible to develop supplemental or alternate versions of the item that provide valid measures of the intended construct for students with the overlapping access need.

In such cases, decision rules must be developed about whether to present the item to students and how to incorporate the item in the total test score.

- Develop an access needs profile for every student, and provide professional development and tools to support decisions about access needs. The concept of an access needs profile will be new to many educators. To help ensure that sound decisions are made about the assignment of access needs, educators will need training. Where possible, tools that help educators make informed decisions should also be developed and made available to teachers in all schools.
- Use test delivery systems that can interpret student access profiles, flexibly activate embedded access tools, and selectively present default, supplemental, and alternate item content based on each student's need. The power of accessible test design requires a sophisticated test delivery engine that is able to read and interpret student access profiles and that can use a profile to customize the delivery interface and the presentation of supplemental and alternate accessibility information based on the profile. Early adoption by NECAP demonstrates that tailored delivery is possible for a limited number of access needs. Careful design and development must build on this early adoption to meet a wider variety of access needs.

Notes

- 1) Robert Mislavy, et al., “On the Roles of External Knowledge Representations in Assessment Design,” *Journal of Technology, Learning, and Assessment* 8(2) (2010). Retrieved January 21, 2010 from <http://www.jtla.org>.
- 2) Center for Universal Design (CUD), “About UD: Universal Design Principles” (1997) http://www.design.ncsu.edu/cud/about_ud/udprincipleshtmlformat.html (accessed February 13, 2009). Archived at <http://www.webcitation.org/5eZBa9RhJ>.
- 3) D. Rose and A. Meyer, “Universal Design for Learning,” *Journal of Special Education Technology* 15 (1) (2000): 66-70.
- 4) NECAP states are: New Hampshire, Rhode Island, and Vermont.