Universal Design for Computer-Based Testing (UD-CBT) Guidelines

CAST Project Members

Robert P. Dolan

Kelly S. Burling

David Rose

Rebecca Beck

Elizabeth Murray

Nicole Strangman

Jennifer Jude

Pearson Project Members

Michael Harms

Walter (Denny) Way

Elizabeth Hanna

Amy Nichols

Ellen Strain-Seymour

Pearson and CAST

October 2010 (Revision B)

PEARSON

Using assessment and research to promote learning

Table of Contents

Introduction

The Universal Design for Computer-Based Testing (UD-CBT) guidelines is a systematized representation of the multi-dimensional UD-CBT framework (Harms et al., 2006; Burling et al., 2006) to support test item development and analysis of item designs. These guidelines are organized according to three tiers: test delivery considerations, item content and delivery considerations, and component content and delivery considerations. The component content and delivery considerations tier is further sub-organized according to the various categories of processing students apply during testing; the former two tiers consider the processing categories implicitly. These processing categories, which will be defined in greater detail shortly, were developed from the principles of Universal Design for Learning (UDL, Rose and Meyer, 2002). They provide a logical framework to organize the guidelines within the component categories, and they facilitate the identification of those guidelines relevant to particular student populations, most notably but not exclusively students with disabilities.

In this introduction the UD-CBT framework is outlined and definitions of the terminology used in the framework and the guidelines are provided. Suggested methods for applying the guidelines in a test-development process are discussed with references to test specifications applications, item development and item reviews. The guidelines themselves follow. The first tier contains test delivery considerations for item development. Item content and delivery considerations, which follow in the second tier, are considerations relevant to individual items, but generic to item type and components. The component content and delivery guidelines follow. Implementing the component content and delivery guidelines requires a careful analysis of an item's intended construct and what the item will actually measure given its components. A further distillation of the component content and delivery guidelines is included as checklists in the Appendix on page 162. These checklists can be used to identify sources of construct-irrelevant variance (CIV) in item designs, and point toward design solutions that can remedy them.

Variance: Observed and Intended

The UD-CBT framework provides a foundation for using digital technologies to create tests that more accurately assess students possessing a diverse range of physical, sensory, and cognitive abilities and challenges. The framework centers around understanding item-level observed score variance. Since items require students to have certain skills, knowledge, and abilities to interact with and respond to them, students with different levels of these requisites will perform differently. The UD-CBT framework is built on the assumption that this item-level variance is a result of what the item asks students to do, regardless of what the item developer intends to measure.

An item can be thought of as a pool of potential variance consisting of the potential variance associated with each individual *component* within the item. The Standards for Educational and Psychological Testing (AERA, APA, NCME, 1999) provide the following definition for variance components:

> *"In testing, variances accruing from the separate constituent sources that are assumed to contribute to the overall variance of observed scores. Such variances, estimated by the methods of the analysis of variance, often reflect situation, location, time, test form, rater, and related effects (p.184)."*

Just as an item requires certain skills, knowledge and abilities, every component in an item requires skills, knowledge and abilities. For instance, reading a poem requires such diverse requisites as visual ability including acuity and discrimination, knowledge of the language, syntax, and vocabulary with which the poem is written and analytical skills to identify what is important to read. These are all required before the student makes sense of the poem, or has considered what type of response is required. If this item were multiple-choice, text in the stem and response options present additional challenges similar to the poem, as well as requirements to understand the multiple choice format, to identify how to choose an answer, and the physical act of marking the correct answer on the test form or on a bubble sheet. Each dimension of each component is a potential source of variance.

The potential for variance exists because of differences in student knowledge, skills, and abilities. If the assumption is that a fully representative sample of students is responding, then all potential sources of variance related to an item's components will contribute to item performance differences. However, only the variance related to the item's intended construct contributes to valid interpretations of student performance. There are two ways to restrict the potential variance in an item to increase the proportion of variance related to the intended construct. First, limit the population of test takers to those who have certain requisite skills, abilities and knowledge. For example, restricting functionally blind students from a paper-based test would ensure that basic visual perception does not contribute to differences in student performance. The second means to restrict potential variance is to design items with built-in supports that reduce or remove sources of variance. Both these methods restrict potential variance by removing the impact of construct-irrelevant variance. As Messick (1994) wrote:

"In the threat to validity known as 'construct-irrelevant variance,' the construct is too broad, containing excess reliable variance that is irrelevant to the interpreted construct (p.8)."

A primary challenge for test developers is to clearly define the constructs they intend to measure. Only when the knowledge, skills and abilities relevant to the intended construct are distinguished from those that are irrelevant can developers address the impact of CIV by limiting the population of examinees for whom the results are valid, or by designing items with built-in supports.

Construct definition is the most fundamental step in test development (Haladyna & Downing, 2004), as well as one of the most challenging. Test-level construct definition is the process of determining the purpose of the test and the domain, i.e. the skills, knowledge and abilities the test intends to measure. Item-level construct definition begins with the test specifications and continues as individual items are developed. Millman and Greene (1993) list topics in the test development process that relate to test and item-level construct definition: purpose, population, timing, administration conditions, sources of content, dimensionality of content, domain breadth, distribution of items across content, item format, size of item pool, item psychometric properties, item selection, test format and production, item scoring, and test-level psychometric properties. These categories set the parameters for what is being measured, how it is to be measured, and what the measurement will mean. If a topic, or a dimension of a topic has not been included in the test specifications then it should not appear in an item, or variance related to student performance on it should be controlled. The onus is on item developers to examine the potential methods of measuring constructs, and to choose item content that best corresponds to the topic and the dimensions of knowledge, skills, and abilities defined in the test specifications. This responsibility applies to defining construct for a test, and to the more detailed process of item-level construct definition.

A perspective on the implications of item construct is provided by Bejar et al. (2003) in their discussion of Embretson (1983):

"*Construct representation* [is] *a key aspect of test validity concerned with understanding the cognitive mechanisms related to the item solution and item features that call on these mechanisms.*"

The issue of cognitive mechanisms and the item features that enlist them is central to construct definition. In the UD-CBT framework and guidelines these item features are referred to as item components. Each component "calls on" mechanisms, or requires certain processing capabilities, from students. Bejar, et al. (2003) recognize cognitive mechanisms associated with item components. The UD-CBT framework and guidelines divide these cognitive mechanisms into six categories of processing, as discussed below.

An item does not generally measure only one mechanism, or "chunk" of discrete knowledge. Indeed, each individual component in an item enlists multiple mechanisms and processes, and each mechanism and process introduces a potential source of variance. An item's construct is the whole of what the item is designed to measure; how well the item measures it determines the validity of inferences that can be drawn from student responses to the item. If the item does not sufficiently measure the construct, it suffers from construct underrepresentation. If, however, it represents the construct, but also asks about extraneous facts or requires mechanisms and processes to be used that are not specified in the item construct, the item is clouded by CIV. Alignment between the intended and measured constructs can be achieved by carefully identifying the purpose of measurement, and matching the purpose to the components that call on the relevant processing capabilities. However, while careful matching of purpose and content will address construct representation and minimize CIV, it cannot eradicate CIV. Each component requires complex interactions of which only a subset will be related to the intended construct. Supporting the interactions that are not related to the intended construct reduces their impact on student performance, and their contribution to item variance.

The Role of Digital Technologies

Built-in supports that can be flexibly accessed to support diverse student needs are realizable through the use of new media and digital technologies. Digital formats can incorporate multiple representations (text, video, audio), transform within and across media, and incorporate tools such as highlighters and linked dictionaries. By combining digital media and technology in computer-based testing (CBT), items can be designed with multiple representations, appropriate tools, and appropriate supports incorporated. This results in a more accessible assessment that allows a greater proportion of the student population to participate without post-hoc accommodations that often interfere with intended measures. It also addresses the impact of CIV on all students during the item design process, resulting in more valid and precise score interpretations for all students

The UD-CBT framework and guidelines have been designed to assist developers in identifying the sources of variance present in their items, and design solutions to ameliorate those that result in CIV. The guidelines represent a strategic initial investigation of item components and CBT interface features to identify sources of variance, and to relate them to functional characteristics of student populations and design options that can minimize construct-irrelevant sources of variance. Research developments in variance, assistive technologies (broadly defined), and item design can be incorporated within the UD-CBT framework and allow the guidelines to expand along with knowledge and technological developments.

Elements of the Framework: Components, Processing, and Students

Defining the sources of variance associated with particular item components requires a systematic definition of item components and a systematic analysis of how students interact with item components. Item components were defined by identifying potential constituent parts of CBT items. The list of components is not exhaustive, and assumes that additional components will be defined and analyzed as CBT grows. The current list of components is shown in Table 1.

Table 1. Item Components and Definitions.

CBT Item Components	Definition
Text	Construct relevant terms or concepts in the instructions, stem, or stimulus materials
Images	Photos, static images (artwork, maps, cartoons, etc.), icons (images on interface elements used to represent functionality), symbols (images that are commonly understood to represent a particular concept)
Audio	Independent audio recordings or an audio track accompanying a video or animation
Tables and Graphs	Tables used to organize information, convey structure and relationships. Graphs used to represent data visually
Mathematical and Scientific Notation	Mathematical expressions, scientific expressions, scientific notation, scientific elements, numbers and symbols
Video and Animation	Visual representations that contain action
Response Options	Actions ranging from typing numbers or characters, clicking a box, clicking on a graphic or text, dragging icons or text while responding to item formats including selected response (multiple choice, multiple response, figural response [select part of a figure or graphic], ordered response [order or sequence a list of items in accordance with some rule]), sorting or categorizing problems or ranking items by correctness; constructed responses include typing a numerical answer to a quantitative question and figural responses where the student marks on, assembles, or interacts with a figure (build a circuit, plot points on a grid, correct errors in a passage)
Active Objects/Links	Words or icons that result in an action or take the user to a different location, or pictures with multiple active regions each which take the user to a different location
Constructed Response: Text	Language-based composition ranging from fill-in-the-blank to essays
Constructed Response: Math	Input a response ranging from a single number to complex proofs or displays of work
Multi-stage/Multi-part Items	Multiple actions or responses required within one item. Screen elements or environment changes at each stage of multi-stage items. Multi-part items have a different page for each part.

The analysis of the components was organized by applying elements of UDL, as described in the companion research report, “The Universal Design for Computer-Based Testing Framework: A Structure for Developing Guidelines for Constructing Innovative Computer-Administered Tests.” UDL was the basis for defining the six *categories of processing* students apply when interacting with components, as shown in Table 2. The options are based on differences in students' skills and abilities to interact with their knowledge and novel information. Essentially, the knowledge and information correspond to curricular content or domain knowledge, while skills and abilities refer to cognitive or fluid abilities (Haladyna and Downing, 2004). Thus, UDL captures the broad categories of explicit and latent traits important in most educational testing. By using UDL as a basis for defining the categories of processing students use when they interact with items, these categories are also associated with the relevant traits of educational testing from a pedagogical perspective.

Table 2. Categories of Processing.

Perceptual processing refers to the activities involved in converting and categorizing the many sensations that reach our brain into stable and valid representations – percepts – of the external world. Through such processes we are able to recognize and remember objects and events in the environment in spite of variation in their sensory features (their size, color, location, in different contexts). Individuals differ in their ability to sense and categorize information between different modalities (vision, touch, etc.) and within different aspects of any modality (pitch, loudness, duration).
Linguistic processing refers to the specific perceptual activities involved in recognizing the patterns of auditory, visual, and tactile stimuli (e.g. Braille) that constitute language. Through specialized processes devoted to language, we are able to recognize and remember the elements from which meaning can be derived: vocabulary, syntax, visual word recognition, text structure (letter, play, poem) etc. Individuals vary in their ability to process language and linguistic elements that is separable from the variance associated with their overall perceptual and cognitive capacities.
Cognitive Processing refers to the skills and strategies by which an individual constructs meaning from the elements of perception and language to interact with his or her environment. Such "meaning making" typically involves the connection and comparison of one element (an object, a word) with other elements in memory (prior knowledge), or in the environment (context), etc. Comprehending text, as opposed to merely recognizing its elements, involves cognitive processing. Individuals differ both in the prior knowledge they can bring to bear in making meaning, and in the strategies and skills they have available to construct that meaning.
Motoric Processing refers to the processes through which meaningful patterns of action can be constructed. Such meaningful patterns of action, a complement to the processes of perception, involve many different forms or modalities of expression – from pointing to speaking to writing text. Each of these forms of expression involve complex patterns of motoric activity, from motor planning to actual execution. Individuals differ in their ability to express themselves both within modalities (spelling versus composition) and across modalities (speaking versus writing).

Table 2. Categories of Processing. (continued)

Executive Processing refers to the processes by which an individual sets and maintains goals, devises plans and strategies for reaching those goals, allocates and organizes the effort and mental resources that would be necessary for implementing those strategies, and monitors progress in reaching goals so that plans can be revised or extended as results warrant. Typically executive processes also involve the processes by which goals and tactics are "held in mind" (e.g. working memory) and the processes by which potential tactics are "tried out" in mind with the intent of predicting their outcomes before actual concrete action is taken. Executive functions are emphasized during novel or unstructured tasks, when more than fact retrieval or routine operations are required. An adult who walks in a novel, challenging, or dangerous environment pauses even their gum chewing to focus attention and cautiously plan the next step.
Affective Processing refers to the processes by which an individual evaluates the importance or significance of events, objects, or plans. Beyond mere recognition of an object, affective processing evaluates its "value". An object's value is determined not by properties of the object but by the interaction between the individual's status – their goals, fears, needs, biological states (like hunger) – and the object's properties. While some affective reactions to the environment are fast, hardwired, and "instinctual" research has also shown that there is a hierarchy of structures which, like the other systems, provide higher levels of integration and opportunities to learn affectively – providing a way that emotional "background knowledge" and new self regulation strategies can be brought to bear on our affective responses to the world and our ability to control and manipulate our emotions productively. How an individual performs on a task is markedly influenced by their affect resulting from both task-relevant and task-irrelevant conditions.

The categories of processing define factors that impact the performance of all students. However, the functional characteristics of individuals and disability groups can be associated with particular categories of processing that will most impact their performance. For instance, perceptual processing refers primarily to visual and auditory access abilities, and thus the sources of variance in each category are most relevant to students with visual and auditory disabilities. As another example, linguistic processing is related to visual and language-based semiotics from vocabulary and visual conventions to decoding, syntax, and fluency; thus sources of variance in linguistic processing are most relevant to students with functional limitations in decoding and understanding the basic elements of textual and visual literacy. Table 3 provides an overview of the disability categories that were the focus of the UD-CBT Guidelines project, and the processing categories primarily affected. Details on the derivation of this table are provided in the companion research report, *The Universal Design for Computer-Based Testing Framework: A Structure for Developing Guidelines for Constructing Innovative Computer-Administered Tests*, as well as in the companion interim deliverable report, *Exemplar Students*, delivered September 2006.

Table 3. Disability Categories and Primary Categories Processing Commonly Affected.

Disability Category	Primary Processing Category(ies)
Blind	Perceptual, Visual
Low Vision	Perceptual, Visual
Deaf, Hard of Hearing	Perceptual, Auditory
Learning Disability: Reading/Language	Linguistic
Mild Mental Retardation	Cognitive
Physical Disability	Motoric
Dyspraxia/Dysgraphia	Motoric
Attention Deficit/Hyperactivity Disorder	Executive
Learning Disability: Math	Executive
Autism Spectrum Disorders	Affective
Emotional Disturbance	Affective

Together the item components and processing categories provide a framework for identifying and organizing sources of variance. By reducing items down to their component parts and analyzing these components according to the categories of processing, the identified sources of variance can be described independently of how the component is used and other components in the item. This in turn supports the development and review of items that minimize the interference of construct-irrelevant factors.

Guidelines Structure

The intent of the UD-CBT Guidelines is to assist item developers in identifying design options that can reduce sources of construct-irrelevant variance and to ensure that the item content is aligned with the measurement goals. Even in instances where it is impractical to implement specific design options, the guidelines support systematic identification of the sources of variance to consider when drawing inferences about student performance.

As discussed previously, aligning an item's content to the specified construct(s) is a fundamental step in ensuring an item is measuring what is intended. These guidelines provide general support for this process; proper alignment can only be achieved with knowledge of the item's intended construct(s) and content as dictated by test specifications and/or the test development process. Rather, a primary purpose of the guidelines is to assist developers in identifying sources of variance in item designs. This is not intended to minimize the role of content selection. Item content, including the medium through which it is conveyed, should be selected by finding the content and media that best tap the knowledge, skills, and underlying processes the item intends to measure. To this end, the guidelines can be used in advance of item development to help item writers understand the knowledge, skills, and processes that the different item components require students to employ, guiding them toward using those that correspond to the intended construct. For instance, reviewing the animation component section of the guidelines shows that there are significant perceptual and linguistic (visual syntax) requirements. This may guide an item writer away from animation for an item intended to measure factual recall. However, if the item is intended to measure students' understanding of complex relationships between objects, concepts, or actions, an animation may be an appropriate way to encapsulate and present the information. The developer can then use the guidelines to identify ways to support access to the item content without changing the intended construct, by identifying the construct-irrelevant sources of variance and their related design options.

Content-construct alignment is primarily relevant to item-level considerations. However, CIV can be introduced at the test level, the item level, and at the within-item component level. The UD-CBT Guidelines contain three tiers at which CIV can be addressed: test-level delivery, item-level content and delivery, and the within-item component-level content and delivery. The *test delivery considerations* address issues generic to all items within a test and do not deal with alignment between intended and actual measurement constructs. The *item content and delivery considerations* present generic questions about the alignment between the item content and the intended construct, and address how CIV can be introduced by how aspects of the item function together in the overall item design. The *component content and delivery considerations* help identify specific sources of variance introduced by different item components, along with recommendations on how to reduce or remove their impact on student performance should they be deemed construct-irrelevant.

Test Delivery Considerations

The test delivery considerations should be applied prior to item development. The actions and guidelines outlined at the test level create a template for item writers and designers. This template ensures a consistent look and feel across items in a test and that items function in a consistent, intuitive and expected manner. Because all tests are developed for distinct circumstances with particular measurement interests, some of the steps in this tier of the guidelines suggest the creation of additional guidelines or modification of existing guidelines. This will support individual test development processes (e.g. individual state tests), just as test specifications need to be customized to support the development of individual tests.

Item Content and Delivery Considerations

The item content and delivery considerations provide support for item development that applies across the various item components. They also serve to ensure that the test delivery considerations are incorporated at the level of item design. For instance, consider the test delivery consideration that rules should be developed for how tools and objects should behave throughout a test. The corresponding item content and delivery guideline checks that the operability of tools and objects in the items is consistent with their use throughout the entire test. The item content and delivery guidelines are also focused on the alignment between the item's construct/constructs and the content being used to assess it/them. These guidelines are intended for use during item development and review.

Component-Level Considerations

The component content and delivery considerations are arranged by individual item components and support a detailed analysis of the alignment between actual and intended constructs measurable with that component present in an item. The guidelines are further subdivided according to the relevant categories of processing.. Within each of these categories, specific sources of variance introduced by the item component are described. If these sources of variance are deemed construct-relevant, recommendations are provided on how to reduce or remove their impact on student performance. Supporting the component content and delivery considerations are corresponding checklists in the Appendix on page 162.

Applying the Guidelines

One model for applying the three tiers of the UD-CBT Guidelines to test item design would consist of the following steps:

1. **Evaluate the item design for construct validity**
   This analysis determines whether the appropriate content was selected given the test item specification. The item-level construct should clearly define the knowledge, skills, abilities, and processes the item intends to measure. The content should be evaluated for construct validity through the following two steps:

   1.1. Assess the construct using the item content and delivery considerations.

   1.2. Assess the construct using test-specific item writing guidelines and item review processes. Test specific item writing guidelines should suggest the scope of the construct, the methods and materials that would be used to best teach the constructs, provide exemplars for how the constructs should be measured, define item formats, and identify potential sources of variance (Smisko, Twing, & Denny, 2000).

2. **Evaluate the item design for sources of construct-irrelevant variance**
   This analysis determines whether the item design and chosen interfaces interfere with the construct by adding additional skill or knowledge requirements due to the interaction between the student and the medium.

   2.1. Identify *item components* within the test item design.

   2.2. Consider the *student group* targeted for this item in terms of their functional skills and limitations. Which processing categories are impacted by the targeted student groups? If the impacted processing categories are construct relevant the item will not be valid for student from the targeted group. Is there a different way to measure the construct that would be valid for all students?

   2.3. For each component, determine which processing categories are construct relevant. Within the construct relevant categories, determine the sources of variance that are construct relevant.

   2.4. Evaluate the item design against the *design recommendations* for all the construct irrelevant sources of variance across all processing categories.

   2.5. For each construct-irrelevant source of variance: if the item design already contains remedies suggested in the design recommendations, or other solutions, no further action is needed.

3. **Revise the item design to incorporate relevant design recommendations and re-evaluate using the UD-CBT process.**
   3.1. For all construct irrelevant sources of variance without solutions embedded in the existing item design, consider the design solutions suggested in the UD-CBT Guidelines. The solutions should be considered according to feasibility, cost-effectiveness, and student populations impacted.
4. **Upon completion of design changes identified through UD-CBT the same process should again be applied to ensure that no additional sources of construct-irrelevant variance were introduced.**

The outcome of these steps process would be the identification of elements of the test item design likely to represent sources of construct-irrelevant variance and proposed methods for minimizing that variance. While no design evaluation can guarantee that all potential sources of construct-irrelevant variance are eliminated, it is expected that a systematic, comprehensive evaluation based on research and best-practices will identify a greater number of issues early. This will facilitate the creation of innovative, computer-based assessment items that are accessible to the widest range of students and more completely achieving the goals of universal design.

Test Delivery Considerations

The test delivery considerations apply to the test development process as well as to more specific aspects of test design. In paper-based testing, guidelines are established for how objectives will be measured, including determination of eligible content and format (Smisko, Twing, & Denny, 2000). Additional guidelines apply to the layout of items in a test booklet, the tools that will be available to students, and how students will progress through the test booklet. The UD-CBT test delivery considerations reflect the same concerns as standard recommendations for the test development process (Millman and Greene, 1993), but apply them to the novel aspects of testing delivered via computer. CBT item writers are faced with numerous tools, media, and functionalities that are not available within traditional assessments. Without guidance and boundaries to establish consistency, every item in an assessment could require students to learn a new interface and new functions, a proposition that would introduce significant sources of CIV.

The UD-CBT test delivery considerations are not prescriptive, nor will applying them remove flexibility or individuality from the test development process. Primarily, they address basic accessibility and compatibility issues, and provide guidance on incorporating CBT features into the test development process.

1. Develop a standardized user interface design that provides the following functionality:

1.1. Provide a set of user tools that will be available across content areas, that shall include the following functionality as a minimum:

1.1.1. Ability to cross-out distractors on multiple choice items

1.1.2. Ability to highlight any portion of an item

1.1.3. Ability to erase portions of constructed responses

1.2. Sequential item-to-item navigation (i.e. next item; previous item)

1.3. Interactive "table of contents" view of all items in current section of test that:

1.3.1. Indicates which item/items is/are currently displayed

1.3.2. Indicates which items have been marked by student for review

1.3.3. Provides interface for navigating to any other item

1.4. Interface for moving between stages of multi-stage items

1.5. Interface for marking individual item for later review

1.6. Interface for undoing the last action on constructed response items

1.7. Interface for resetting responses on constructed response items

1.8. A consistent location and layout of tools

2. Ensure interface template is available to all item writers and developers so they have a sense of screen design, layout, and real estate.

3. Develop specifications for the presentation and combination of components in items with multiple components in the stimulus materials.

3.1. Provide a menu or similar representation of all available stimulus materials visible throughout the item so students can navigate through them at will.

4. Develop specifications for the interface and functionality of interactive components.

5. To the extent that students will be able to take notes, draft responses, have access to scrap paper, use graphic organizers, etc. develop specifications for how these additional materials will be accessed, and when and how they will be visible (toggle, minimize, only available as a separate page, available superimposed over material on the screen, etc.)

6. Based upon curriculum standards and common materials and instructional approaches, determine what activities, simulated environments, and reference materials are relevant for assessment.

7. Based on the activities, simulated environments, and reference materials, design a set of tools, a list of appropriate simulation environments, and a list of activities to be used by item writers.

7.1. Examples of tools include dictionaries, ruler, calculator, formula sheets, etc.

7.2. Examples of simulation environments include searching for references in a library, performing a chemistry experiment, or choosing a path through town that requires interacting with various individuals and conversing in a foreign language.

7.3. Examples of activities within the simulation environments include measuring chemicals (manipulating a scale, pouring, scooping, etc.), mixing chemicals (adjusting the flame on a Bunsen burner, stirring), performing a computer-based search of a library's holdings (typing, clicking, navigating with a mouse).

8. **Use the specifications developed from the previous guidelines to create a style sheet for item writers and developers/programmers.**
9. **Adhere to current best-practices for accessible user interface design. The most relevant include:**

   9.1. World Wide Web Consortium (W3C) Web Content Accessibility Guidelines and User Agent Accessibility Guidelines (http://www.w3.org/WAI/guid-tech.html)

   9.2. Checklist for Web Content Accessibility (http://www.w3.org/TR/WAI-WEBCONTENT/full-checklist.html). Excerpts from this checklist:

   - Provide a text equivalent for every non-text element
   - Ensure that all information conveyed with color is available without it
   - Identify changes in the natural language of a document's text and any text equivalents
   - Organize documents so they are readable without style sheets
   - Ensure that equivalents for dynamic content are updated concomitant with the content
   - Avoid screen flickering
   - Use the cleanest and simplest language appropriate for a site's content (apply to construct-irrelevant language such as test directions and consider for item directions, questions, and distractors)
   - Provide redundant text links for each active region of a server-side image map
   - Provide client-side image maps instead of server-side image maps except where the regions cannot be defined with an available geometric shape
   - Title each frame to facilitate frame identification and navigation

   9.3. Trace Research and Development Center at the University of Wisconsin Application Software Design Guidelines (http://trace.wisc.edu/docs/software_guidelines/toc.htm). General design guidelines synopsis:

   - Use system tools whenever possible
   - Maintain consistent, predictable layout & behavior and adhere to system standards/style guides
   - Provide keyboard access to all dialogs, menus, and tools

- Design software to minimize the skills and abilities needed to operate it (unless those skills and abilities are construct-relevant to the assessment)
- Be sure software cooperates with (or at least does not break) special access features in the OS and third party software
- Use and open systems approach

9.4. IMS Global Learning Consortium Guidelines for Developing Accessible Learning Applications (http://www.imsproject.org/accessibility/accv1p0/imsacc_guidev1p0.html).

9.5. Section 508 of the U.S. Rehabilitation Act (http://www.section508.gov/index.cfm?FuseAction=Content&ID=12)

10. Ensure operating system-level accessibility features are available or functionally replicated.

10.1. Features that reduce potential sensory barriers such as visual display and auditory controls are often provided by the host operating system. These features must not be disabled. If they are, equivalent features must be made available by the test administration software.

10.2. Features that reduce potential motoric barriers such as keyboarding and pointing device controls are often provided by the host operating system. These features must not be disabled. If they are, equivalent features must be made available by the test administration software.

Item Content and Delivery Considerations

The item content and delivery considerations are designed to support a comparison of an item's intended and actual constructs, prior to analyzing the specific processing requirements of each item component. Component-level analysis alone is insufficient for ensuring the item is valid for its intended purpose and audience. The component analysis addresses what can be done to make an item accessible once it is included in an item. The item content and delivery considerations encourage item developers to consider what components are most appropriate for their measurement purposes, and the implications of using multiple components simultaneously.

The relevance of item content to the item's construct, and the clarity with which the content is presented, directly impact item validity. If the information is misaligned to the construct or is not clearly contextualized and presented, the functionality of the item components hardly matters because item validity has already been compromised. The relationship between item content and construct validity discussed in the UD-CBT Guidelines is based on an accepted definition of validity (Messick, 1993), traditional item writing guidelines (Haladyna, 1999, Haladyna, Downing & Rodriguez, 2002) and research into sources of CIV in assessment items (Leighton & Gokiert, 2005).

Holistic analysis of the item in terms of its relationship to other items on the test, and its relationship to test specifications also reinforces concepts expressed in the test delivery considerations. For instance, guideline 2 in the item content and delivery considerations is directly related to guidelines 6,7, and 8 in the test delivery considerations.

1. **Relevant (item-level): the item measures the construct(s) it intends to measure without extraneous content or tasks.**
    1.1. Does the item clearly address knowledge, skills, and/or abilities identified in the test specifications?
    1.2. Are the content and tasks of the item clearly related to the objectives the item is supposed to measure?

2. **Relevant (component-level): the components chosen to measure the intended construct(s) are the most appropriate components for the purpose.**
    - 2.1. Does the construct suggest a particular medium? For example, use of text for the spelling section of an English/Language Arts assessment, and use of images for identifying cellular structures in a biology test.
    - 2.2. If the construct can be measured equally well using any number of different components, is it feasible to include all of them or a subset of them? If not, has the most accessible component been chosen?
3. **Representative: item elements correspond to materials and/or environments used in the classroom.**
    - 3.1. Do the content and structure of the item align with how teachers and experts consider quality instructional methods?
    - 3.2. Does the item look like something students will have seen or used in the classroom?
4. **Realistic: there is an unambiguous relationship between media or virtual environment and its real-world counterpart.**
    - 4.1. If media is used to represent an actual process or event is it sufficiently realistic to be easily identifiable?
    - 4.2. Could stimuli be mistaken for something else, or be too removed from an actual representation that matching the media to what it is intended to represent introduces construct-irrelevant cognitive load?
5. **Synergistic: item elements complement one another in conveying meaning.**
    - 5.1. Do multiple elements stimulate the same processing category simultaneously, do they compete?
    - 5.2. Are there multiple simultaneous visual or auditory stimuli?
6. **Clear and unambiguous: item intent and expected response are conveyed clearly and with contextualization.**
    - 6.1. Do the instructions clearly convey the scope and intent of the item?
    - 6.2. Are the steps necessary to reach the end, and how to proceed through them clear?
    - 6.3. Is the context sufficiently defined?

7. **Free of bias: item is sensitive to the full population of test takers.**
    - 7.1. Is the item sensitive to cultural, socio-economic, gender, age, language, disability and regional issues?
    - 7.2. Will prior knowledge unfairly advantage one group over another?
    - 7.3. If post field-testing, does the DIF analysis indicate bias?
8. **Consistent across items: item format, tools, and operability are the same throughout a test; students do not need to learn something new in order to respond to an item**
    - 8.1. Are all elements presented in a manner consistent with other items (e.g., items appearing underneath text, test navigation along the bottom of the screen, blue underlined text indicates a link)?
    - 8.2. Are all tools located in the same space in the design?
    - 8.3. Are all tools and functionalities accessed in the same way as in other items throughout the test?
9. **Appropriate time and task load: the time required to view and interact with item elements is appropriate to the intended difficulty and level of inquiry of the item; the impact of the item on student's time or energy to complete the rest of the test has been considered**
    - 9.1. Is the length of any multi-media elements appropriate for the difficulty of the item and the level of inquiry (e.g. recall versus problem solving)?
    - 9.2. Is the time required to interact with an item appropriate for the difficulty and level of inquiry of the item?
    - 9.3. Do design features, such as multiple screens, increase task load inappropriately (e.g. multiple screens taxing working memory when the measurement focus is drawing inferences between text)?
    - 9.4. Is the time or task load of an item going to negatively impact the amount of time or energy a student will have to spend on other items?

Component Content and Delivery Considerations

The component content and delivery considerations are arranged by individual item components. For each of the components there is a general discussion of the issues surrounding the use of the component in terms of accessibility. This is followed by the actual considerations/guidelines, arranged by the relevant categories of processing.

Note that the individual component sections were written to stand on their own. No assumptions are made as to whether users of these guidelines would also refer to other components. Furthermore, there was concern that too much cross-referencing would compromise the usability of the guidelines. As a result, there is redundancy of information between component sections.

Text

Text is a ubiquitous form of communication that is predominant in educational materials. However, extracting meaning from text requires many different perceptual, linguistic, cognitive, executive, affective, and even motoric (e.g. eye movement) processes. Since individual students vary considerably in all of those processes, the use of text potentially introduces a great deal of item-level variance – which may or may not be construct relevant. While text is not inherently accessible for many individuals, when it is presented using appropriate digital technologies many of its barriers can be significantly reduced.

Universal design of text is nearly impossible in a print environment but comparatively easy in a digital one. With upcoming federal regulations such as the National Instructional Materials Accessibility Standard (NIMAS) for the presentation of curricular materials, the availability of highly structured and flexible digital text will become commonplace (NIMAS, 2006). With NIMAS and other digital formats such as the Digital Accessible Information System (DAISY Consortium, 2006), predictable and automated transformations of text into alternative formats can be achieved. These alternate formats include read-aloud (through text-to-speech), playback of synchronized recorded audio, Braille, and signing avatars. The basic flexibility of these digital formats (especially when combined with editorial markup schemas for supporting cognitive and executive processes that are currently under development) creates an excellent foundation for providing the alternatives that are necessary to reduce construct-irrelevant variance associated with the use of text.

Traditional assessments use text to convey the directions, navigation, stimulus, prompts, and item content. Nearly all states (46) allow the directions to be read aloud and, nearly as many allow questions to be read aloud, but only in certain circumstances (31) or only in certain circumstances with implications for scoring (13) (Clapper, Morse, Lazarus, Thompson, & Thurlow, 2005). Read-aloud accommodations for item questions continue to be controversial. The extent to which construct-relevant content is conveyed in the question and the extent to which reading that information to students modifies the construct is often unclear. For example, does a reading test require that students be able to interact with the printed word, or that the students can successfully make meaning from words? The National Accessible Reading Assessment Projects (NARAP) is currently exploring this question. NARAP is a collaboration of two projects funded by the U.S. Department of Education, Office of Special Education Programs (OSEP), to conduct research on how to make large-scale assessments of reading proficiency more accessible for students who have disabilities that affect reading (NARAP, 2006). Central to their work is exploring the relationship between the definition of reading and the appropriate forms of assessment. Interacting with text also affects student performance in subject area tests. For example, reading and understanding directions as well as item content is necessary in math (particularly word problems), social studies, and science.

UDL advocates providing access to educational materials, such as test content, without construct modification. However, in the present context there are policy implications and significant costs involved in implementing such accessibility. The decision about whether to include text-to-speech capabilities, whether to make tests accessible to screen readers, and the implications for scoring must continue to be made on a case-by-case, state-by-state basis. Involving test developers, reading experts, parents, students, teachers, and other stakeholders in the decision process can help ensure that all needs are considered and that all constituents are aware of the final decision and its rationale. The focus of such decision-making processes should be on construct definition, and the construct relevance of the text in each item.

> The Text Component Checklist on page 162 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Text Component Considerations – Overview

Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination
Linguistic Processing:	· English Language Proficiency · Vocabulary Knowledge · Syntactic Skills · Decoding and Fluency Skills · Knowledge of Text Structure
Cognitive Processing:	· Background Knowledge · Comprehension Strategies · Categorical and Conceptual Skills · Concentration and Attention
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. Visual Ability

1.1. Providing access to text for blind students requires multiple options. Some students use Braille displays, while others use screen reading software, and others may use a combination. Deaf blind students will only be able to use Braille. Providing instruction and practice in how applications or devices will interact with the test system is necessary prior to actual testing.

1.1.1. Screen Readers and Self-Voicing Output: Most screen reading software uses Microsoft Active Accessibility (MSAA) protocols (Microsoft Corporation, 2006); these can be accessed at http://msdn.microsoft.com/library/default.asp?url=/library/en-us/msaa/msaastart_9w2t.asp. Screen readers generally handle text in single columns best, and often have difficulty when multiple hypertext markup language (HTML) frames are used.

1.1.2. Applications which provide internal text-to-speech rendering of text-based files are known as self-voicing text-readers. Capabilities such as reviewing text, reviewing individual words or sections, and checking the spelling of words that are mispronounced or homonyms need to be included in self-voicing applications. Including text-to-speech capabilities in a test application removes the need to provide assistive technology during testing, but the developer must ensure use and functionality that match common screen reading applications and allow test takers to practice using the self-voicing technology.

1.2. Refreshable Braille displays generally present only one line of text at a time, and those with less than 80 cells provide only part of a line at a time. Most portable displays have 40 cells, but several 85 cell displays are available. Some displays have keys that can be programmed for additional functionality, such as mouse key commands. Many are designed to be used in conjunction with screen reading software. Most Braille displays are designed to work with the Windows operating system, applications should use the aforementioned MSAA protocols to ensure compatibility. Braille displays are unable to present images, tables, icons, etc. but they can present alt-text (alternative text) and long descriptions with d-links associated with images, tables, icons, etc. in well-formed HTML.

1.3. HTML tips for accessible text from Building Accessible Websites (Clark, 2002):

1.3.1. Declare the character encoding of all documents.

1.3.2. Use structural rather than presentational text elements, including headings.

1.3.3. When revising old documents and producing new documents, use `<abbr></abbr>` and `<acronym></acronym>` to mark up abbreviations and acronyms that would be unfamiliar to a typical visitor.

1.3.4. Mark up block quotations with `<blockquote></blockquote>`, but do not use that element for other purposes.

2. Visual Acuity

2.1. All fonts used should allow examinees to adjust size and/or be amenable to the use of cascading style sheets (CSS).

2.1.1. As a default, 12 point fonts are considered standard for paper, fonts between 12 and 18 are considered enlarged, and 18 point fonts are considered large print (Allman, 2004).

2.1.2. The impact of font size adjustments on the item's layout should be considered, and any design changes necessary to maintain the general look and feel of the item should be automatically invoked when corresponding font sizes are chosen. Text should be allowed to reflow or rewrap when the font size changes.

3. Visual Discrimination

3.1. Black text on white or pastel background generally has the highest readability for most students. For low-vision students reverse contrast should be made available either through operating system-level features or directly in the test administration software.

3.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

3.3. Sans-serif fonts (e.g. Verdana, Arial) should be used as they generally have higher readability on-screen than serif fonts (e.g. Times, Palatino).

Linguistic Processing

Considerations for reducing barriers due to differences in:

4. English Language Proficiency

4.1. Alternates to English language text should be offered for students who are English language learners (ELL), American Sign Language (ASL) for hard-of-hearing, and Braille for blind and deaf/blind students. Consideration must be made of both the student's native or preferred language and the language of instruction. Providing students with choice of language on an item-by-item basis, such as through dual-language presentation, may be a the best option provided students have demonstrated success with such conditions during practice sessions. In addition, language translation dictionaries can be provided online for look-up of individual words or phrases.

4.2. Alternatives to text should be considered when other media, such as images, video, animation, etc. could also convey necessary information, perhaps in ways that are easier for students to comprehend.

5. Vocabulary Knowledge

5.1. Providing vocabulary support through definitions for individual words can be handled by rollovers/mouseovers or links to dictionary and/or thesaurus tools. Dictionaries and thesauruses should be available in multiple languages and provide translation to English. A dictionary and thesaurus tool should provide a signing avatar to translate words into ASL for deaf students. There are several commercial providers of signing avatars such as Vcom3D (Vcom3D Inc., 2006), as well as several available for free download on the web; see http://www.signingbooks.org/animations/sign_language_animations.htm for a description of available technologies, uses and how avatars are created (Pragma-NL, 2002).

5.1.1. Rollovers or mouseovers that provide definitions that appear next to a student's mouse pointer when he or she holds the computer mouse over a predefined word allow the developer to control the words for which definitions are provided. However, they can clutter screen real-estate, may not be equally apparent or accessible to all students, and can cause confusion for screen readers. Rollovers/mouseovers and pop-ups are not accessible to many screen readers unless they can be locked in place so a screen reader can navigate to and read them (California Community Colleges, 2000).

5.1.2. Providing access to an online dictionary through a tool link requires students to exhibit initiative to find definitions. Differences in student motivation and initiative could impact performance and introduce CIV. However, providing a dictionary instead of rollovers/mouseovers allows students to find definitions for words other than those specified by the test developers, which can alleviate CIV due to differences in vocabulary knowledge. Providing a spell checker with the dictionary tool will help students find words that they might spell incorrectly; such a tool is currently built into most web search engines. This would be especially important for looking up words that are heard through audio rather than read.

5.1.3. Use appropriate markup language to facilitate pronunciation for screen readers or self-voicing reference tools. See W3C-WAI (1999) guidelines for examples and techniques (http://www.w3.org/TR/WAI-WEBCONTENT/#Guidelines).

6. Syntactic Skills

6.1. Syntactic skills affect students' abilities to derive meaning from text that is conveyed through the organization of words, textual elements, and structure. Grammar aids and simplified syntax can be used to make text more accessible. Juicy Studio (2006) (http://juicystudio.com/services/readability.php) has an online tool that generates readability statistics for text passages or web pages, according to several research-based algorithms.

7. Decoding and Fluency Skills

7.1. Text-to-speech technology can support decoding and fluency, thereby potentially reducing or removing barriers to comprehension.

7.1.1. Decoding refers to the ability to figure out how to read, pronounce and derive meaning from unknown words by using prior knowledge of letters, sounds, phonemes, and word patterns. Text-to-speech can be used to allow individual words to be spoken aloud (or signed into ASL) or spoken and defined with a talking dictionary. As with vocabulary supports, implementing text-to-speech capabilities for all words or only for selected words in the item text requires careful consideration so as to not modify constructs.

7.1.2. Fluency is the ability to read text with appropriate speed, accuracy, and prosody, all of which depend upon decoding ability. Text-to-speech allows students to simultaneously read and hear text spoken aloud and thus focus on comprehension without the of decoding and fluent reading. Synchronous highlighting is often used in conjunction with text-to-speech to support comprehension by focusing readers' attention.

8. Knowledge of Text Structure

8.1. Explicit identification of text structure (i.e. letter, speech, short story) provides contextual clues that might otherwise confound students.

8.2. Graphic organizers provide a visual representation of text content and structure. They can be used as advance organizers to help activate and organize background knowledge, as a means to organize thoughts and information during reading, or as post-organizers after encountering a text (Strangman & Hall, 2002). In each instance, the graphic organizer could accompany the text in the item layout, or be made available through an external link.

8.2.1. When presented in advance of a text passage, a graphic organizer introduces the structure, prepares readers for the content and assists in comprehension.

8.2.2. During reading, a graphic organizer can be used as a note taking template, simultaneously reasserting the text structure. Whether the organizer is entirely blank, or contains some level of detail needs to be determined by item developers with respect to construct modification.

8.2.3. Graphic organizers can leverage the text structure to support specific comprehension strategies. For example, a graphic organizer that enables students to take notes in a time line format while reading a history item stimulus passage will reinforce a summarization strategy, since recalling sequence is one way to formulate a summary. It is common instructional practice to leverage text structure in this way as a support for using metacognitive strategies such as summarization, questioning, predicting, self-monitoring, evaluating, etc.

Cognitive Processing

Considerations for reducing barriers due to differences in:

9. Background Knowledge

9.1. Links to relevant prior knowledge can be provided to complement reading of text passages. Links could be embedded in the text by creating words or phrases that link to definitions, brief descriptions, and/or images. Links could also appear as a menu at the beginning or alongside a text passage. Topic knowledge has been shown to impact what and how much is comprehended (Stahl, 1991, 1989; Hirsch, 2003). The extent to which links to material provided during testing, material which is succinct enough to be used within a testing session, can improve performance of students who lacked relevant background knowledge is unknown. Judicious use of such supports is recommended as there is also the potential for confusing or slowing students down, negatively impacting their performance.

10. Comprehension Strategies

10.1. Text passages that do not fit on a single screen can affect comprehension. Scrolling has been shown to have a negative impact on comprehension relative to paging through text but the relationship is confounded by examinee facility with computers (Higgins, Russell, Hoffman, 2005). In CBTs that are to coexist with paper versions, mirroring the presentation of text across the two modes minimizes the impact of mode on performance (Pommerich, 2004), but is often impractical and negates technology's ability to reformat text.

10.2. Reminders of comprehension strategies can be incorporated into test instructions, attached to item directions, or embedded in text passages. Research using hypertext with embedded supports has shown positive effects on comprehension from providing information about reading strategies, (Anderson, Horney, & Blair, 1999-2001) inserting comprehension (Tobias, 1987, 1988) or meta-cognitive questions (Dalton, Pisha, Eagleton, Coyne, & Deysher, 2002; Salomon, Globerson, & Guterman, 1989) and providing training in metacognition and reading strategies (Dalton et al., 2002; McNamara, Levinstein, & Boonthum, 2004).

11. Categorical and Conceptual Skills

11.1. Advance organizers, concept maps, and other graphic organizers can support categorization and conceptual understanding, in addition to providing information about text structure. Saye (2001) describes the use of graphic organizers and other scaffolds to increase comprehension in a hypermedia environment. See Section 8.2 above for details.

12. Concentration and Attention

12.1. Prompting could be embedded throughout a test, or within particular items by having attentional reminders appear on screen or as an auditory alert according to predetermined time intervals. Brooks, Todd, Tofflemoyer, and Horne (2003) have shown increases in performance due to prompting.

12.2. Breaking text into smaller sections with associated items permits students to answer questions while the relevant information is fresh and the relevant text is in front of them. Myles & Simpson (2001); Safran (2002); and Unok-Marks et al. (2003) have found comprehension and performance benefits for students with autism spectrum disorders with prompting and breaking text and long tasks into stages; such an approach would likely be successful for students with attention deficit/hyperactivity disorder as well.

Executive Processing

Considerations for reducing barriers due to differences in:

13. Goal Setting Ability

13.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for reading text in the item more salient or evident. For some students, the fact that the actual purpose or goal for reading text in an item is often implicit rather than explicit is an impediment to reading strategically – and thus to comprehension, recall, selectivity, problem-solving (e.g. Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

13.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for reading is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for reading. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

14. Goal Maintenance and Adjustment

14.1. Particularly for reading long passages of text, some students will have difficulty in maintaining a consistent goal or purpose for their reading (some, for example, will find themselves continually interrupted by sub-tasks like decoding, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks of reading are not construct-relevant (e.g. decoding, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

14.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the text, or the articulation of longer passages into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer passages and items (Wood, 1988; 2002).

14.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

15. Progress Monitoring

15.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals so s in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

16. Working Memory

16.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

17. Self-regulation

17.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

18. Intrinsic Task-specific Motivation

18.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing reading comprehension or problem solving that requires reading text, it is advantageous to provide two (or more) different passages which differ in content (e.g baseball versus ballet) but which maintain equivalent levels of difficulty in vocabulary, syntax, etc. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

18.2. In earlier guidelines the importance of providing alternatives to traditional printed text has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to text by students with different histories and abilities in text environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

19. Extrinsic Incentives

19.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on reading comprehension, the uniform use of extrinsic rewards by all students who are reading text represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

19.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in text reading. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

20. Test Conditions

20.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

20.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

20.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

20.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

20.5. Items should be reviewed for age appropriate content.

Images

In the UD-CBT Guidelines images refer to photos, icons, illustrations, maps and diagrams. As with print literacy, the first level of visual literacy is basic identification of the elements (identifying the salient features of images) and the second level is comprehension, which in the case of images involves critical thinking (Thibault & Walbert, 2003). Reading images incorporates observation, critique, deconstruction, and discernment of point of view or bias. The first level of visual literacy is primarily perceptual, and presents a challenge for students who are blind or have low-vision. Understanding visual material is challenging for students who have difficulty integrating pieces of information into a conceptual understanding, and students who do not have strategies for analyzing and interpreting information. Certain types of images, such as charts and diagrams, can exacerbate this situation due to unfamiliar use or placement of text.

Images are used in assessment to convey information, or to complement information presented through an alternate medium, such as text with illustrations. Images are often a more convenient way to convey information, but are rarely used as the sole means for transmitting information. Maps, cartoons, and even photo content can be appended with text descriptions, which can be conveyed to blind and low vision examinees through alt-text and long descriptions with d-links; useful tips on creating meaningful alt-text can be found at http://www.pantos.org/atw/35534.html (All Things Web, 2006). The challenge is creating descriptions that sufficiently convey the image without giving clues to the item answer.

Traditional item writing guidelines suggest that images should only be used when they are necessary to answer the question (Haladyna, 1999). Extraneous or gratuitous images can be confusing, distracting and negatively impact student performance. However, for some students images enhance understanding. The benefits of illustrations for understanding a text varies across students, task, type of pictures, and learning material, with some complex interactions occurring between these factors (Filippatou & Pumfrey, 1996; Schnotz & Bannert, 2003). In our own student interviews, students consistently asserted that images were helping in reducing cognitive load associated with reading, outweighing any potential confusion from the images. In the literature, some studies indicate that poor readers are more likely to be helped by illustrations than are good readers, while other studies indicate that they are less helped by illustrations (see Filippatou & Pumfrey, 1996; Levie & Lentz, 1982; Schnotz & Bannert, 2003). The usefulness of pictures also appears to interact with domain knowledge (Hegarty & Just, 1989; Kuntz, Drewniak, & Schott, 1989). Students with less prior knowledge may benefit more from pictures (Kozma, 1991). Kozma (1991) concludes:

> *"…for some learners, the use of pictures, in addition to text, may provide information needed to map mental representations derived from the text onto mental representations of the real world. (p.188)."*

Gambrell & Jawitz (1993) demonstrated that a dual strategy of looking at illustrations and making a mental image led to significantly better recall of story elements in fourth graders than did only looking at illustrations.

Williams (1993) lists six circumstances where graphics can provide information that are difficult to derive from text and therefore are more effective than text: to describe or identify, to concretize the abstract, to convey spatial information, to provide a meaningful context for unfamiliar information, to help the reader solve a problem, to convey procedural information, and to scaffold memory. These instances may provide guidance for when and how to use images in item development, or at least guidance for further research on images in assessment. Williams also notes that visuals can increase the speed of performance with procedures, ensure a standardized mental image, and increase the ability to remember concrete events or concepts. These assertions have implications for instructing examinees in the procedures necessary to interact with an item and for standardizing examinee interpretation of material. However, he also indicates that readers minimally process images presented in conjunction with text unless explicitly directed to do so. These findings agree with a recent study involving tracking of students' eye movements as they read. In this study students were found to increase their viewing of images when images were placed into text columns and when the images were explicitly cued within the text; this was especially true for struggling readers who increased their viewing of image twofold as a function of layout (Dolan, 2006).

In addition to refining the role of images in traditional assessment items, it is important to realize that images have a increasing role in education due to technology and media in the classroom. There is a contemporary call for a broader definition of literacy which includes drawing and visual aspects of literacy, in part because the contemporary world demands multimedia literacy (Bousted & Ozturk, 2004; Kendrick & McKay, 2004; Pope Edwards & Mayo Willis, 2000). Moore (2000) lists six modes of visual learning that could also be applied in assessment to generate novel items. His focus is on photography, but the six modes could also be applied to other visual media: exploring (e.g. using pictures to investigate details of objects), recording (arranging photographs to depict an event), expressing (using photographs to record feelings or to talk about them), motivating (using photographs to document student achievement and motivate students), communicating (using photographs to create a story), and imagining (using photographs to inspire imaginative thinking, for example passing around a photo and creating a "story in the round", asking students to describe what is happening in photos).

> The Images Component Checklist on page 163 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Images Component Considerations – Overview

Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination · Color Perception · Shape Recognition
Cognitive Processing:	· Visual Processing · Using Graphic Conventions · Using Iconic Conventions · Visual Syntax Fluency · Background Knowledge · Comprehension Strategies · Interaction Strategies · Planning and Organizing Skills · Concentration and Attention
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. **Visual Ability**

   1.1. All images should be accompanied by alt-text and long descriptions with d-links. Guidance and examples of these can be found at http://www.ncddr.org/du/researchexchange/v02n01/design.html (Southwest Educational Development Laboratory, 2004), http://ncam.wgbh.org/cdrom/guideline/guideline1.html (NCAM 2003), and http://www.w3.org/TR/WCAG10-CORE-TECHS/#text-equivalent (W3C-WAI, 1999).

   1.1.1. Care must be taken to ensure that text descriptions do not lead provide clues to the correct answer or provide information not available to other test-takers.

   1.1.2. Text equivalents or audio descriptions should be created by subject matter experts who have knowledge/experience with test-takers with visual impairments.

   1.2. Tactile graphics can be produced by many Braille printers, but they are most appropriate for simple drawings or line renderings. According to the American Printing House for the Blind (Allan at al., 2003):

> *"Most maps, charts, graphs, and diagrams can be made tactual if the test publisher will allow some editing. Editing could involve eliminating shading used solely for visual effect, reducing the number of distracters, providing two or three charts to present the same information as a complex print chart, using descriptions to supplement or replace graphics, or using symbols and words with a key to provide information," (p.14).*

   1.2.1. Provide tactile images using Swell or puff paper, which can go through printers and copiers; it is available from Repro-Tronics Inc. (http://www.repro-tronics.com). Radiant heat causes black printed areas to swell. The Tiger Software Suite from View Plus Technology (http://www.viewplustech.com) used in combination with tactile printers (such as Emprint Haptic Color Braille Embosser, also from View Plus) allows tactile renderings of full screen shots or particular areas on the screen.

   1.2.2. Tactile graphics should be accompanied by descriptive text, multiple images may be required. According to IMS (2002):

> *"Tactile graphics may be most effective when accompanied by text that introduces students to the conventions used and guides them in exploring the graphic. Less information can be conveyed in a tactile graphic than in a visual one because the sense of touch has*

lower resolution than vision does, so it may take more than one tactile graphic to capture the information completely in a complex visual diagram."

1.3. Additional renderings of images can be created using audio and haptic devices. For more information see http://ncam.wgbh.org/cdrom/guideline/guideline7.html. Haptic mouse devices are available from Immersion Corporation (San Jose, CA, USA) and Control Advancements (Kitchener, Ontario, Canada). These devices can allow users to explore images tactilely. Obviously, students should be familiar with such devices before employing them during assessment. .

1.4. Dynamic images will either be animations, or require screen refreshing to display changes.

1.4.1. For animations refer to the Video/Animation Component section below.

1.4.2. For screen refreshing highlight areas on the screen with new information for visual users. For screen readers and Braille devices ensure that there is coding to direct the device's focus to the new information without re-reading the entire page. Information on how to handle this with Document Object Models (DOM) is available at: http://juicystudio.com/article/dom-screen-readers.php.

1.4.3. Prevent, or allow users to disable, automatic refresh of animations or movies; screen readers restart reading a page when an element on the page is automatically refreshed.

1.5. Three-dimensional manipulatives can be provided to accompany items for which tactile graphics are not sufficient, or when students are not versed in the conventions of tactile representations of 3-D materials or images.

1.6. For images embedded using HTML see Building Accessible Websites (Clark, 2002), Chapter 6 for an extensive discussion of coding options for simple and complex images including cutaway drawings, graduated rating scales, genealogical charts and flowcharts, outlines, and other hierarchies. http://joeclark.org/book/sashay/serialization/Chapter06.html

1.7. Comics can be handled using alt-text and long descriptions with d-links. In XML, Jason McIntosh's ComicsML set of document type definitions are available at jmac.org/projects/comics_ml/.

2. Visual Acuity

2.1. When possible images should represented using scaleable vector graphics (SVG) or equivalent, since images can be scaled without degradation. W3C SVG specifications are available at: http://www.w3.org/TR/SVG/

2.2. 3X zoom renders most graphics un-interpretable due to pixelation. If vector graphics is not an option, all bitmap images should be provided at multiple resolutions.

2.3. Allow images to be printed separately from the rest of the page on which it appears. These can then be used to create enlarged images for low vision students.

3. Visual Discrimination

3.1. Dark images on white or pastel background generally has the highest readability for most students. For low-vision students reverse contrast should be made available either through operating system-level features or directly in the test administration software.

3.2. If possible, complex images should be simplified to provide an interpretable high-contrast image or line drawing. Text description should accompany simplified images to ensure the full meaning is conveyed.

3.3. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

4. Color Perception

4.1. Avoid using hue differences as the sole means of conveying information. Hue differences combined with luminance and/or texture differences are fine.

4.2. Avoid common color blindness combinations; provide monochromatic option, user specified color combinations. To check visibility to individuals with various types of color blindness see: http://www.vischeck.com/vischeck/vischeckURL.php. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 JuicyStudio, 2006).

5. **Shape Recognition**
   - 5.1. Avoid using shape as the sole means of conveying information, i.e. shapes as icons.
   - 5.2. Provide alt-text and long descriptions with d-links of the shapes in an image if they are central to the content.
   - 5.3. Provide a tactile representation of shapes for blind and low-vision students.
   - 5.4. When areas in a single screen layout have different functions, use different shapes, colors, or other identifying features to distinguish between them.

Cognitive Processing

Considerations for reducing barriers due to differences in:

6. **Visual Processing**
   - 6.1. Highlight critical features in images; provide instructions to direct attention to relevant areas or features of images.
     - 6.1.1. Text in images is inaccessible to screen readers unless it is made part of the alt-text and long descriptions with d-links.
     - 6.1.2. Text in images should be reproduced separately from the image if it is in a fixed font, if it is not oriented left-justified and horizontal, and for students who have difficulty with visual search.
7. **Using Graphic Conventions**
   - 7.1. Do not assume knowledge of graphic conventions. Instead provide descriptions and literal depictions.
8. **Using Iconic Conventions**
   - 8.1. Avoid using icons as the sole means of conveying content or directions.
     - 8.1.1. Provide a legend for icons used throughout a test, accessible at any time through a link or toggle page.
     - 8.1.2. Provide customizable icons, i.e. let examinees choose icons to represent tools or commands at the beginning of the test.

8.1.3. Use rollovers or mouseovers to provide definitions of icons (rollovers and mouseovers are not inherently accessible to screen readers unless they can be locked and a cursor directed to them).

9. Visual Syntax Fluency

9.1. Highlight critical relationships between features in images, or between images.

9.2. Provide instructions about the proper sequence of images if sequence is relevant to comprehension.

9.2.1. Use highlighting or focus shift to draw attention to images in sequence if sequence is relevant to comprehension.

10. Background Knowledge

10.1. Links to relevant prior knowledge can be provided to complement understanding of images. Links could be embedded in the image by creating client side image maps with links to definitions, brief descriptions, and or images (image maps should include alt-text for links). Links could also appear as a menu at the beginning or alongside a text passage. Topic knowledge has been shown to impact what and how much is comprehended (Stahl, 1991, 1989; Hirsch, 2003). The extent to which links to material provided during testing can improve the performance of students who lack relevant background knowledge is unknown. Judicious use of such supports is recommended as there is also the potential for confusing or slowing students down, negatively impacting their performance.

11. Comprehension Strategies

11.1. Reminders of viewing and interpretation strategies can be incorporated into test instructions, attached to item directions, or programmed to appear at timed intervals throughout the item interaction.

12. Interaction Strategies

12.1. Students may wish to draw directly on stimuli as they work through the problem. Rather than requiring students to redraw images on scratch paper and thus introduce potential sources of challenge (e.g. inclusion of all detail, matching of scale), students should be provided with scratch paper that already contains representations of the stimuli or be allowed to print hardcopies of on-screen stimuli.

13. Planning and Organizing Skills

13.1. Advance organizers, concept maps, and other graphic organizers can support categorization and conceptual understanding, in addition to providing information about structure. They can be used as advance organizers to help activate and organize background knowledge, as a means to organize thoughts and information during viewing, or as post-organizers after interacting with an image. In each instance, the graphic organizer could appear in the item layout, or be made available through a link.

13.1.1. When presented in advance a graphic organizer introduces the structure, prepares viewers for the content and assists in comprehension.

13.1.2. During viewing, a graphic organizer can be used as a note taking template, simultaneously reasserting the structure and relationships in the image. Whether the organizer is entirely blank, or contains some level of detail needs to be determined by item developers with respect to construct modification.

13.1.3. Graphic organizers used after viewing images can be used to summarize and clarify content and structure. They can also be presented blank, or nearly blank, and used by viewers to summarize and clarify their comprehension.

14. Concentration and Attention

14.1. Prompting could be embedded throughout a test, or within particular items by having attentional reminders appear on screen or as an auditory alert according to predetermined time intervals. Brooks, Todd, Tofflemoyer, and Horne (2003) have shown increases in performance due to prompting.

14.2. Simplified images can make important relationships more clear by removing "background noise."

Executive Processing

Considerations for reducing barriers due to differences in:

15. Goal Setting Ability

15.1. Images present very different problems from text and/or audio – they are primarily "simultaneous" rather than "sequential" – our vision takes in the whole image at the same time that it is sequentially focusing on parts. Because of that simultaneity, the role of setting goals for investigating images is often underestimated (Yarbus, 1967, Gregory, 2000) but students differ considerably in their ability to investigate images systematically and often do not set a purpose for viewing that will guide their search effectively. To make the goal for investigating images more salient, provide explicit instructions and/or initial prompts – "Here's the purpose for looking at the image. Here is what to look for." Whether to use such prompts in an item will depend on whether goal-setting is construct relevant.

15.2. In items where prompting on the actual goal for looking would invalidate an item, prompts that are more "generic" can be used, scaffolding students to set a purpose for looking, to make a plan for what they will look for, a strategy for remembering what they have found, and so forth.

16. 2. Goal Maintenance and Adjustment

16.1. Many students will have difficulty in maintaining a consistent goal or purpose for investigating an image; they will be drawn to salient details rather than driven by their goal or objective. Within a CBT environment it is possible to construct and display images with many different scaffolds embedded to help students in maintaining their focus on reaching goals – e.g. sequential display of images so that only certain parts are displayed at a time but in an order that guides the student to look at relevant information in a goal-driven sequence, sequential highlighting so that certain elements are highlighted in an order that facilitates recognition and comprehension. At any or all transition points it is possible to embed prompts to remind student what the goal for viewing is.

16.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" and structural supports (summarizing statements, emphasis of transitions, etc.) can be provided to scaffold students in their viewing of an image. Additionally, comprehension checks can be embedded within the item in order to assist students in focusing on the "main ideas" etc. in the image that are relevant to task demands.

17. Progress Monitoring

17.1. In a CBT environment it is possible to represent progress visually (a concurrent timeline, or a rolling outline) or auditorily through verbal prompts and markers that make the structure of the item more explicit. However, this does not sufficiently scaffold them in monitoring their progress toward effective problem-solving, finding answers to a question, and so on. For that purpose, it is possible to add additional scaffolds for the processes involved in using active listening as a step in problem solving: a checklist that indicates how many steps have been achieved in a multi-step problem, a graphic display of progress, embedded comprehension checks that provide feedback ("you have found three causes of the Civil War, four are needed") (Deno, 1999, Noonan and Miller, 1995)

18. Working Memory

18.1. Exploration of an image places less demands on working memory than listening, for example (Osaka et al. 2004, Jambaque, et. al., 1993) because the image typically is re-viewable. The issue to address in addressing working memory differences tends to reside more in the ability to select relevant information, remember it, and construct responses from that information. External memory aids can be provided for checking or accumulating information from graphics or images – a notebook (either paper or on screen), a checklist, organizer, etc. – so that the working memory load is reduced during the process of constructing an answer, etc.

Affective Processing

Considerations for reducing barriers due to differences in:

19. Self-regulation

19.1. General supports for self-regulation while viewing an image – the ability to sustain motivation over extended periods, to respond effectively to threats and challenges in investigating an image, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within more difficult items (a complex diagram, or a comparison among several graphics, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

20. Intrinsic Task-specific Motivation

20.1. Where specific content is not construct relevant (e.g. where the item is specifically assessing image processing skills, for example, rather than specific content knowledge) providing alternative images will reduce threats to validity that arise from difference in interests and background knowledge. For example, in assessing the ability to use graphic displays, it may be advantageous to provide two (or more) different graphic displays which differ only in the topic addressed (e.g relationship between batting averages of older and younger baseball players versus the relationship between income and age). Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

20.2. Especially with complex diagrams, or especially for students who have visual difficulties, the initial presentation of an image or graphic may elicit significant affective reaction that impedes effective problem solving and comprehension. The importance of techniques for providing alternatives to a single fixed presentation has been emphasized above for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to images by students with different histories and abilities in text environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like text description or animation of the graphic, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

21. Extrinsic Incentives

21.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

21.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in listening. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

22. Test Conditions

22.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT focused on images as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

22.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

22.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

22.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

22.5. Items should be reviewed for age appropriate content.

Audio

Tasks of the audio listener include establishing a frame of reference, interacting with the materials, and assimilating the message (Brothers, 1971). Listening and reading differ with respect to the types of cues available, decoding demands, permanence of the stimuli, and the rate of information. While there appears to be some overlap between reading and listening comprehension, and some inferences can be made about students' reading comprehension based on their listening comprehension performance, the two types of processing are distinct (Carlisle& Felbinger, 1991; Sinatra, 1990). Students have different strategies for understanding and remembering print and speech (Carlisle & Felbinger, 1991). Listening comprehension is influenced by learner (level of intelligence, age, and academic achievement; Brothers, 1971) as well as outside variables such as the purpose for listening, and concentration. Brothers (1971) suggests offering a preview of material and/or auditory cues to help establish a frame of reference, and facilitating interaction by pausing the content, or allowing it to be paused, to offer opportunity for reflection and response.

Auditory perception must also be considered for students who are deaf or hard-of-hearing. However, deriving meaning from the audio content presents linguistic and cognitive challenges similar to those of text. There are also motoric challenges related to navigating the audio file.

In the UD-CBT Guidelines and framework, audio refers to stand-alone audio recordings as well as the audio tracks that might accompany video or animation. Since the implementation considerations were the same for both cases, they are covered together in this section (rather than incorporating audio tracks into the same section as video/animation).

> The Audio Component Checklist on page 164 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Audio Component Considerations – Overview

Perceptual Processing:	· Hearing Ability · Auditory Threshold · Auditory Discrimination and Processing
Linguistic Processing:	· English Language Proficiency · Vocabulary Knowledge · Syntactic Skills · Prosodic Recognition · Using Idiomatic Expressions
Cognitive Processing:	· Background Knowledge · Comprehension Strategies · Planning and Organizing Skills · Concentration and Attention
Motoric Processing:	· Object Manipulation and Navigation Abilities
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. **Hearing Ability**

    1.1. Closed captioning accompanying a video or animation ensures that audio components are accessible to students who are deaf or hard-of-hearing. It may be incorporated into multi-media as open (always visible to all users) or closed (visible only to users who turn on the audio description). Synchronized Multimedia Integration Language (SMIL) and Synchronized Accessible Media Interchange (SAMI) formats support the inclusion of captioning. SMIL is played by Quick Time Player, Real Player, and Oratrix GRiNS Player. SAMI is played by Windows Media Player. Flash animation technology allows native captions to be created within an animation. MAGpie (Media Access Generator) can be used to create captions for SMIL and SAMI. More information about SMIL, SAMI and MAGpie, including techniques and code can be found at http://ncam.wgbh.org/cdrom/guideline/guideline2.html and http://ncam.wgbh.org/richmedia/ .

2. **Auditory Threshold**

    2.1. Volume should be adjustable throughout the test for alerts and test sounds. A volume control should accompany each video/animation and stand-alone audio component.

3. **Auditory Discrimination and Processing**

    3.1. Audio with emphasized discriminants could be incorporated as an option for students who are hard-of-hearing. The rate of delivery should be adjustable, and the audio should be reviewable.

Linguistic Processing

Considerations for reducing barriers due to differences in:

4. **English Language Proficiency**

    4.1. Alternates to English language audio should be offered for ESL students, American Sign Language (ASL) and captioning for hard-of-hearing.

5. **Vocabulary Knowledge**

    5.1. Vocabulary support through definitions for individual words can be provided by dictionary and thesaurus links.

5.2. Providing access to an online dictionary through a tool link requires students to exhibit initiative to find definitions. Differences in student motivation and initiative could impact performance and introduce CIV. However, providing a dictionary instead of rollovers or mouseovers allows students to find definitions for words other than those specified by the test developers, which can alleviate CIV due to differences in vocabulary knowledge. Providing a spell checker with the dictionary tool will help students find words that they might spell incorrectly; such a tool is currently built into most web search engines. This would be especially important for looking up words that are heard through audio rather than read.

5.2.1. Use appropriate markup language to facilitate pronunciation. See W3C guidelines for examples and techniques (http://www.w3.org/TR/WAI-WEBCONTENT/#Guidelines).

6. Syntactic Skills

6.1. Syntactic skills affect students' abilities to derive meaning from speech that is conveyed through the organization of words, textual elements, and structure. Grammar aids and simplified syntax can be used to make text more accessible.

6.2. Simultaneous and synchronized presentation of text and audio in text-to-speech applications is likely to increase student comprehension.

7. Prosodic Recognition

7.1. Alternative audio tracks with various prosodic emphasis could be offered (native language, natural language, regional dialect variations).

7.2. Offering students choice of voices in text-to-speech applications improves their opportunity to find a voice they understand.

7.3. Allowing students to adjust rate of speech in text-to-speech application allows them to maximize comprehension. Students should be allowed to change rate on-the-fly as they may wish to slow voices down during particularly challenging passages.

8. Using Idiomatic Expressions

8.1. Alternatives to or definitions of idiomatic expressions can be provided (see Background Knowledge below).

Cognitive Processing

Considerations for reducing barriers due to differences in:

9. Background Knowledge

9.1. Links to relevant prior knowledge could be provided alongside a video/animation. Topic knowledge has been shown to impact what and how much is comprehended (Stahl, 1991, 1989; Hirsch, 2003). The extent to which links to material provided during testing, material which is succinct enough to be used within a testing session, can improve performance of students who lacked relevant background knowledge is unknown. Judicious use of such supports is recommended as there is also the potential for confusing or slowing students down, negatively impacting their performance.

10. Comprehension Strategies

10.1. Reminders of listening and interpretation strategies can be incorporated into test instructions, attached to item directions, or programmed to appear at timed intervals throughout the item interaction.

10.2. Navigation controls (pause, fast forward, reverse, replay, and search) permit students to determine their own pace and revisit information on an as-needed basis.

11. Planning and Organizing Skills

11.1. Advance organizers, concept maps, and other graphic organizers can support categorization and conceptual understanding, and provide information about structure. They can be presented in advance of content to help activate and organize background knowledge, as a means to organize thoughts and information during viewing, or as post-organizers after students interact with the media. In each instance, the graphic organizer could appear in the item layout, or be made available through a link.

11.1.1. When presented in advance a graphic organizer introduces the structure, prepares viewers for the content and assists in comprehension.

11.1.2. During listening, a graphic organizer can be used as a note taking template, simultaneously reasserting the structure and relationships in the image. Whether the organizer is entirely blank, or contains some level of detail needs to be determined by item developers with respect to construct modification.

11.1.3. Graphic organizers used after listening can be used to summarize and clarify content and structure. They can also be presented blank, or nearly blank, and used by viewers to summarize and clarify their comprehension.

12. Concentration and Attention

12.1. Prompting could be embedded throughout a test, or within particular items by having attentional reminders appear on screen or as an auditory alert according to predetermined time intervals. Brooks, Todd, Tofflemoyer, and Horne (2003) have shown increases in performance due to prompting.

12.2. Simplified content can make important relationships more clear by removing "background noise."

12.3. Increased segmentation provides opportunities to reflect and comprehend what has been heard.

12.4. Navigation control allows self-pacing and review.

Motoric Processing

Considerations for reducing barriers due to differences in:

13. Object Manipulation and Navigation Abilities

13.1. Provide keyboard alternatives for all on screen navigation and actions.

13.2. Respect operating system defaults to ensure assistive device and software compatibility.

Executive Processing

Considerations for reducing barriers due to differences in:

14. Goal Setting Ability

14.1. To make the goal for listening more salient, provide explicit instructions and/or initial prompts – "Here's the purpose for listening. Here is what to listen for." Preferably these prompts should occur in alternative modalities, like text. For some students, the fact that the actual purpose or goal for listening in an item is often implicit rather than explicit is an impediment to listening – and thus to comprehension, recall, selectivity, problem-solving based on that listening (Handel, 1993). Whether to use such prompts in an item will depend on whether goal-setting is construct relevant.

14.2. Where prompting on the actual goal for listening would invalidate an item, prompts that are more "generic" can be used, scaffolding students to set a purpose for listening, to make a plan for what they will listen for, a strategy for memory aids, and so forth. Such supports can guide students to focus their efforts through appropriate goals for listening (Underwood, 1989, Byrnes, 1989).

15. Goal Maintenance and Adjustment

15.1. Many students will have difficulty in maintaining a consistent goal or purpose for their listening; they will seem distractible or passively inattentive. Providing an option for breaking the listening passage into shorter segments will help some and others will be greatly aided by the ability to change the rate of speech (see above and Tallal et al.). In all cases, providing a structural outline (as appropriate to the construct being measured) or graphic organizer of the listening passage will provide re-entry points for listeners.

15.2. Some students will be distracted by difficulties at lower levels in the processing of speech or sounds (English language learners, students with limited vocabularies, students with auditory processing difficulties) and will lose the top-level goals for purposeful listening. Where such supports can be provided without invalidating the construct measurement, all of the accommodations mentioned under previous sections can provide the low level scaffolding that will allow students to maintain higher-level goals.

15.3. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" and structural supports (summarizing statements, emphasis of transitions, etc.) can be provided to scaffold students in their listening. Additionally, comprehension checks can be embedded within the item in order to assist individuals in maintaining focus on the "main ideas," etc. that are timely enough to help in overall task orientation.

16. Progress Monitoring

16.1. One of the most difficult challenges in a listening item is to track one's progress toward completion because working memory resources are already challenged by the demands of active listening. In a CBT environment it is possible to represent progress visually (a concurrent timeline, or a rolling outline) or auditorily through verbal prompts and markers that make the structure of the item more explicit. However, this does not sufficiently scaffold them in monitoring their progress toward effective problem-solving, finding answers to a question, and so on. For that purpose, it is possible to add additional scaffolds for the processes involved in using active listening as a step in problem solving: a checklist that indicates how many steps have been achieved in a multi-step problem, a graphic display of progress, embedded comprehension checks that provide feedback ("you have found three causes of the Civil War, four are needed") (Deno, 1999, Noonan and Miller, 1995)

17. Working Memory

17.1. Active listening places great demands on working memory (Osaka et al. 2004, Jambaque, et. al., 1993). When the ability to attend, comprehend, and/or retain information in an auditory mode are not essential to the construct being measured, it is possible to supplement auditory information with external memory aids – a text transcript or caption, a concept map or chart, a diagram, that can provide an alternative means of representation or organizational support. Also, when the item calls for constructing a response from auditory elements, or with auditory elements, it may be possible to augment internal resources with external aids built into the CBT – a visual checklist, organizer, or template, or auditory directions and prompts. (for a wider view of listening comprehension and its components see Byrnes, H. (1984), Noonan and Miller, 1995, Handel 1993).

Affective Processing

Considerations for reducing barriers due to differences in:

18. Self-regulation

18.1. General supports for self-regulation while listening – the ability to sustain motivation over extended periods, to respond effectively to threats and challenges in a listening exercize, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long listening passage, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

19. Intrinsic Task-specific Motivation

19.1. Where specific content is not construct relevant (e.g. where the item is specifically assessing auditory listening skills, for example, rather than specific content knowledge) providing alternative content for the listening exercise will reduce content-specific threats to validity that arise from difference in interests and background knowledge. For example, in assessing listening comprehension or problem solving that requires listening, it is advantageous to provide two (or more) different passages which differ in their content focus (e.g speeches about politics or sports figures) but which maintain equivalent levels of difficulty in vocabulary, syntax, length, etc. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

19.2. In earlier guidelines the importance of providing alternatives to a fixed presentation (in loudness, language, concepts, etc.) has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to text by students with different histories and abilities in text environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like an outline or pre-loading of vocabulary, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

20. Extrinsic Incentives

20.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on listening comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

20.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in listening. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

21. Test Conditions

21.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT focused on listening as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

21.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

21.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

21.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

21.5. Items should be reviewed for age appropriate content.

Tables and Graphs

Tables and graphs are often used to represent or summarize information in science and mathematics. They are also found in social studies, and occasionally reading when it is associated with curricular areas outside of English/Language Arts. Tables and graphs are visual representations, and, as such, raise accessibility issues for students with visual disabilities. However, there are ways to make them more accessible and to present the information they contain in a non-visual manner.

The Tables and Graphs Component Checklist on page 165 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Tables and Graphs Component Considerations – Overview

Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination · Color Perception
Cognitive Processing:	· Background Knowledge · Comprehension Strategies · Interaction Strategies · Planning and Organizing Skills · Concentration and Attention
Motoric Processing:	· Object Manipulation and Navigation Abilities
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. **Visual Ability**

   1.1. Avoid use of structural HTML for visual formatting (W3C), and do not use tables unless the information can be linearized and still make sense (see 1.4). There is some disagreement about the extent to which this guideline is an absolute. See http://joeclark.org/book/sashay/serialization/Chapter10.html for a discussion, and code examples.

   1.2. For graphs and tables meant to convey data, it is very difficult to construct sufficient descriptions. The reason data, a text based element, is put into tables and graphs is to render the data more meaningful. Reverse translating the tables and graphs back into data removes the meaningfulness. A long description can be used to summarize the main points, but this risks modifying the construct being measured. A more viable option is to use the html title attribute or long descriptions with d-links to describe the function and design of the table or graph, and let students explore data points within it using tabbed navigation. PopChart [D] by Corda (corda.com/d/), can automatically summarizes numerical charts and graphs, producing a long description and a d-link.

   1.3. When using HTML to create tables:

      1.3.1. Identify all row and column headers, and data cells. Nested tables are particularly difficult for screen readers and Braille devices.

      1.3.2. Tables for numeric data are complex to code and to make accessible to screen readers and Braille devices. See http://joeclark.org/book/sashay/serialization/Chapter10.html for a discussion and code examples.

      1.3.3. Use HTML to identify groups of rows and columns, and to express complex data relationships. See: http://www.w3.org/TR/WAI-WEBCONTENT/ and http://ncam.wgbh.org/cdrom/guideline/guideline4.html

      1.3.4. Present the table information in a linear text-only format so it is accessible. Headers must be associated with each relevant cell of data, and a table summary **must** be provided. See http://ncam.wgbh.org/cdrom/guideline/guideline4.html.

      1.3.5. Use Lynx, a text only web browser, or a similar technology to test table accessibility. Lynx View and Lynx-Me are Lynx emulators.

      1.3.6. Test accessibility using a self-voicing browser (if applicable).

1.3.7. Test table organization for tabbed navigation (essentially how screen readers, Braille devices, and single switch devices navigate).

1.4. Static tables and graphs can be rendered in scalable vector graphics (SVG) or similar technology with long descriptions. SVG supports text descriptions, stylesheets, DOM2, assistive technologies and input devices. Another possibility, Macromedia Flash, has accessibility features for data grids. See: http://www.macromedia.com/resources/accessibility/flash8/

1.5. Provide pre-recorded audio for static tables and graphs using the conventions for spoken mathematics (see mathematical and scientific notation: numbers and symbols). Provide complete text-based descriptions of static tables and graphs.

1.6. For dynamic tables and graphs:

1.6.1. Provide text-based summaries

1.6.2. Dynamic tables and graphs will either be animations, or require screen refreshing to display changes.

1.6.3. For animations refer to the Video/Animation Component section below.

1.6.4. For screen refreshing, highlight areas on the screen with new information for visual users, and for screen readers and Braille devices ensure that there is coding to direct the device's focus to the new information without re-reading the entire page. Information on how to handle this with Document Object Models (DOM) is available at: http://juicystudio.com/article/dom-screen-readers.php. Prevent, or allow users to disable, automatic refresh of animations or movies; screen readers restart reading a page when an element on the page is automatically refreshed.

1.7. Provide tactile renderings of tables and graphs. Swell paper, which can go through printers and copiers, is available from http://www.repro-tronics.com. Radiant heat causes black printed areas to swell. The Tiger Software Suite from View Plus Technology (http://www.viewplustech.com) used in combination with tactile printers (such as Emprint Haptic Color Braille Embosser, also from View Plus) allows tactile renderings of full screen shots or particular areas on the screen.

1.8. Audio accessible graphing calculators can be used to input numeric data in tables, or for graphs. There are many commercially available; a version developed by the Science Accessibility Project at Oregon State University can be seen at http://dots.physics.orst.edu/calculator.html.

1.9. Additional renderings of graphs include using audio and haptic devices. For more information see http://ncam.wgbh.org/cdrom/guideline/guideline7.html. Haptic mouse devices are available from Immersion Corporation (San Jose, CA, USA) and Control Advancements (Kitchener, Ontario, Canada). These devices can allow users to explore images tactilely. Obviously, students should be familiar with such devices before employing them during assessment.

2. Visual Acuity

2.1. The impact of zoom on the orientation, layout, and legibility of tables and graphs should be tested.

2.2. Allow students to manipulate the size and thickness of lines and fonts in tables and graphs.

2.3. Use SVG, Flash, or similar technologies to permit resizing.

3. Visual Discrimination

3.1. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

3.2. Allow user to control fonts used in tables and graphs.

4. Color Perception

4.1. Avoid using hue differences as the sole means of conveying information. Hue differences combined with luminance and/or texture differences are fine.

4.2. Avoid common color blindness combinations; provide monochromatic option, user specified color combinations.

Cognitive Processing

Considerations for reducing barriers due to differences in:

5. Background Knowledge

5.1. Links to relevant prior knowledge could be provided alongside a table or graph. Topic knowledge has been shown to impact what and how much is comprehended (Stahl, 1991, 1989; Hirsch, 2003). The extent to which links to material provided during testing, material which is succinct enough to be used within a testing session, can improve performance of students who lacked relevant background knowledge is unknown. Judicious use of such supports is recommended as there is also the potential for confusing or slowing students down, negatively impacting their performance.

6. Comprehension Strategies

6.1. Reminders of interpretation strategies can be incorporated into test instructions, attached to item directions, or programmed to appear at timed intervals throughout the item interaction.

7. Interaction Strategies

7.1. Students may wish to draw directly on stimuli as they work through the problem. Rather than requiring students to redraw tables and graphs on scratch paper and thus introduce potential sources of challenge (e.g. inclusion of all information in the proper place), students should be provided with scratch paper that already contains representations of the stimuli or be allowed to print hardcopies of on-screen stimuli.

8. Planning and Organizing Skills

8.1. Advance organizers, concept maps, and other graphic organizers can support categorization and conceptual understanding, and provide information about structure. They can be presented in advance of content to help activate and organize background knowledge, as a means to organize thoughts and information during viewing, or as post-organizers after students interact with the media. In each instance, the graphic organizer could appear in the item layout, or be made available through a link.

8.1.1. When presented in advance a graphic organizer introduces the structure, prepares viewers for the content and assists in comprehension.

8.1.2. During viewing, a graphic organizer can be used as a note taking template, simultaneously reasserting the structure and relationships in the table or graph. Whether the organizer is entirely blank, or contains some level of detail needs to be determined by item developers with respect to construct modification.

8.1.3. Graphic organizers used after viewing content can be used to summarize and clarify content and structure. They can also be presented blank, or nearly blank, and used by viewers to summarize and clarify their comprehension.

9. Concentration and Attention

9.1. Prompting could be embedded throughout a test, or within particular items by having attentional reminders appear on screen or as an auditory alert according to predetermined time intervals. Brooks, Todd, Tofflemoyer, and Horne (2003) have shown increases in performance due to prompting.

9.2. Simplified content can make important relationships more clear by removing extraneous information.

Motoric Processing

Considerations for reducing barriers due to differences in:

10. Object Manipulation and Navigation Abilities

10.1. Provide keyboard alternatives for all on screen navigation and actions.

10.2. Respect operating system defaults to ensure assistive device and software compatibility.

Executive Processing

Considerations for reducing barriers due to differences in:

11. Goal Setting Ability

11.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

11.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

12. Goal Maintenance and Adjustment

12.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

12.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

12.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

13. Progress Monitoring

13.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

14. Working Memory

14.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

15. Self-regulation

15.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

16. Intrinsic Task-specific Motivation

16.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

16.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

17. Extrinsic Incentives

17.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

17.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

18. Test Conditions

18.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

18.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

18.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

18.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

18.5. Items should be reviewed for age appropriate content.

Mathematical and Scientific Notation: Numbers and Symbols

Mathematical notation is widely used in mathematics, the physical sciences, economics, statistics, and engineering. Numbers, operators, equations, functions and variables are all part of mathematical notation. It is a language used to precisely convey concepts in mathematics. Mathematical expressions are combinations of numbers and symbols that can be evaluated according to the rules and conventions of mathematical syntax.

Particular challenges to using mathematics in hypermedia environments result from the interaction between screen readers and mathematical expressions. Mathematics expressed as text cannot always be interpreted by screen readers, and mathematics expressed as images cannot be read by screen readers without accompanying text descriptions. These challenges effect perceptual processing, which is the most basic level of accessibility. Perceptual processing will almost always be construct-irrelevant in K-12 assessment, so the design solutions for perceptual access are of particular importance.

When math numbers and symbols are used in items to measure subjects outside of mathematics, it is imperative to determine whether the math is construct relevant before incorporating supports for understanding. Due to the subject's complexity it is important to evaluate what an item contains and what it requires students to do to determine construct relevance. Supports for understanding and computation can be incorporated for construct-irrelevant content.

> The Mathematical and Scientific Notation: Numbers and Symbols Component Checklist on page 166 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Mathematical and Scientific Notation: Numbers and Symbols Component Considerations – Overview	
Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination
Linguistic Processing:	· To Create Accessible Mathematical Syntax · Mathematical Fluency
Cognitive Processing:	· To Create Accessible Contexts · Calculations · Complex Expressions
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. Visual Ability

1.1. When possible mathematical expressions that are graphical in nature should be represented using scaleable vector graphics (SVG) or an equivalent, since images can be scaled without degradation. Additional description for screen readers can be provided using appropriate language (NCAM guidelines for spoken mathematics, Mathspeak for Nemeth Code (http://www.rit.edu/~easi/easisem/talkmath.htm). SVG supports text descriptions, stylesheets, DOM2, assistive technologies and input devices

1.2. LaTEX (http://www.maths.tcd.ie/~dwilkins/LaTeXPrimer/) and MathML (www.w3c.org/Math) are markup languages that provide some support to math in a hypermedia environment.

1.3. Many web browsers do not currently support SVG or MathML. LaTEX is supported by some browsers, such as Mosaic.

1.4. Embedding text and audio files, gets tricky with complex expressions because each part needs to be accessible separately as well wholly. Concatenated speech is awkward for complex expressions.

1.5. ChemML makes notation, results displays, and data for chemistry accessible to screen readers, it needs to be embedded in XML (or XHTML). CML is a version of ChemML, see http://www.xml-cml.org/information/position.html. ChemML interoperates with other mark-up languages and XML protocols: XHTML for text and images; SVG for line diagrams, graphs, reaction schemes, phase diagrams, etc.; PlotML for graphs; MathML for equations; XLink for hypermedia (including atom-spectralPeak assignments, reaction mapping); RDF and Dublin Core for meta-data; and XML Schemas for numeric and other data types.

2. **Visual Acuity**

2.1. All fonts used for numbers or mathematical and scientific expressions should allow examinees to adjust size and fonts. If choice is not an option, size should be relative and not fixed.

2.2. When possible mathematical expressions that are graphical in nature should be represented using SVG or an equivalent, , since images can be scaled without degradation.

3. **Visual Discrimination**

3.1. Black or dark content on white or pastel background generally has the highest readability for most students. For low-vision students, reverse contrast should be made available either through operating system-level features or directly in the test administration software.

3.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

Linguistic Processing

Considerations for reducing barriers due to differences in:

4. To Create Accessible Mathematical Syntax

4.1. Mathematical syntax can be made more accessible by highlighting the order of operations in expressions and providing definitions of mathematical symbols. Syntax will most often be supported when math is used outside of math assessments. However if items on a math test are specifically interested in problem solving and conceptual understanding, it might be appropriate to support access to mathematical syntax.

5. Mathematical Fluency

5.1. Simplified numbers can replace more complex numbers, while retaining the item's concept when the intent of the item is not to measure facility with numbers (Ostad, 1998; Cumming & Elkins, 1999; Geary, 1993, 2004; Hanich, Jordan, Kaplan, & Dick, 2001). For example, in an item about the ratio of the areas of inscribed geometric figures, the values for the areas of the inner and outer figures are presented to 3 decimal places. The concept of the ratio of areas of inscribed figures could just as easily be tested with whole numbers.

Cognitive Processing

Considerations for reducing barriers due to differences in:

6. To Create Accessible Contexts

6.1. Math items that are situated in a context may reference background knowledge that is not part of the construct being measured. Providing links to background knowledge allows all examinees understand the context.

7. Calculations

7.1. Providing on-screen calculators makes the content of items the focus of measurement, not the underlying calculations. If calculation is the construct being measured, access to the calculator should be barred. Some students may prefer stand alone calculators and/or scrap paper. Replacing complex numbers with simplified numbers can also make calculations less difficult (Torbeyns, Verschaffel, & Ghesquiere, 2004; Ellington, 2003; Thompson & Sproule, 2000).

7.2. The Audio-Accessible Graphing Calculator (http://dots.physics.orst.edu/calculator.html) is a self-voicing Windows application that has been under development and testing for some time by the Science Access Project at Oregon State University. It includes the capabilities to:

- Compute and display visually either of two functions, their sum, or their difference.
- Display the above as an audio tone plot.
- Permit piece-by-piece audio browsing.
- Print the above to any Windows printer including the Tiger tactile graphics embosser.
- Be used as a universally usable on-screen scientific calculator.
- Be used as a powerful expression evaluator.
- Input tabulated data for display.
- Compute statistical functions for tabulated data.

7.3. Braille 'n Speak, a Braille calculator that can do standard math and generate tactile graphs (http://www.dinf.org/csun_99/session0113.html). Additional software is available for students needing the power of a financial calculator and graphing calculator functionality is available for the blind using Graphit in conjunction with any Braille embosser.

7.4. Triangle is an application for blind computer users which allows them to read, write and manipulate scientific text, do scientific computations, and read graphs and figures. (http://www.december.com/cmc/mag/1998/feb/gartri.html)

8. Complex Expressions

8.1. Providing pictorial representations can make complex expressions more conceptually accessible (Fuchs, Compton, Fuchs, Paulson, Bryant, & Hamlett, 2005).

8.2. Blind and low-vision students need to be able to interact with complex expressions, including parsing by sections.

8.2.1. AsTeR is an application that reads Latex notation, creates audio of the mathematical expression, and allows the audio to be navigated. (http://www.cs.cornell.edu/Info/People/raman/aster/aster-toplevel.html)

8.2.2. HP EzMath is a notation for embedding mathematical expressions in web pages, based on how expressions are read aloud. It will soon be available as open source at www.sourceforge.net.

8.3. IBM TechExplorer (http://www.software.ibm.com/techexplorer/) is a plug-in for Navigator and Internet Explorer, an IE 5.5 XML Behavior, and an ActiveX control for applications like Microsoft PowerPoint and Word. TechExplorer enables the display of Tex Latex and MathML documents and the publishing of interactive scientific material on the Web. Version 3.1 includes full support for MathML 2.0, augmented Latex display, options for scripting/programming and a web equation editor. It has a Macintosh version, Mathematica connectivity, and Internet Explorer MathML Behavior (IMS Design Guidelines: http://www.imsglobal.org/accessibility/accwpv0p6/imsacc_wpv0p6.html#1273818)."

8.4. Additional math extensions and programming information: A search of www.sourceforge.net using the terms "math application" returned over 50 hits. Some were games and activities, but most related to add-ins, programming, or libraries.

Additional resources for mathematical and scientific images:

- Table of Mathematical Symbols: http://en.wikipedia.org/wiki/Table_of_mathematical_symbols
- Tables of Physical Constants: http://en.wikipedia.org/wiki/Physical_constant
- GIF (Graphics Interchange Format) and PNG (Portable Network Graphics) Images for Math Symbols: http://us.metamath.org/symbols/symbols.html
- Purdue University TAEVIS center provides access to 4,000 images that are ready to be printed and transferred to thermal paper to create tactile graphics. They also provide a manual for creating new images. Their graphics are primarily college level.: http://www.taevisonline.purdue.edu/Tactile_Diagrams.html

Executive Processing

Considerations for reducing barriers due to differences in:

9. Goal Setting Ability

9.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

9.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

10. Goal Maintenance and Adjustment

10.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

10.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

10.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

11. Progress Monitoring

11.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

12. Working Memory

12.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

13. Self-regulation

13.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

14. Intrinsic Task-specific Motivation

14.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

14.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

15. Extrinsic Incentives

15.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

15.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

16. Test Conditions

16.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

16.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

16.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

16.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

16.5. Items should be reviewed for age appropriate content.

Video/Animation

Video incorporates pictures and words, as with books, but is distinctive in several ways: the linguistic information is oral, the two types of information are presented simultaneously, and all content is transient. While readers can adapt their reading to suit the complexity of the text (change their reading speed, skip or re-read difficult content), a video viewer has no such opportunities.

Scwhan & Riempp (2004) suggest that one of the primary advantages of media such as video is their ability to optimize how viewers experience the information ("optimize the experiencing conditions"). Video is customizable in the sense that design decisions (choice of camera lens, use of multiple cameras, post-production arranging and assembling, etc.) can be used to facilitate cognitive processing of viewers (Schwan & Riempp, 2004). However, most mass media have to design based on the "average user" and in this situation the design decisions might be restrictive for many viewers.

Interactive video is a specialized form media that offers a choice of what and how to view (some control over pace and sequence). Interactive video helps improve the match between presentation and individual by turning part of the process of shaping the presentation over to the user (Schwan & Riempp, 2004). However, managing these interactive features also increases cognitive load (Lowe 2004).

Studies with college age students have shown mixed reactions to video used in place of text. In the studies, however, there was no attempt to use video to convey information most appropriate to the medium, it was simply used as an alternate to text. Based on interviews about the study strategies students used, it appeared as if some of the difficulty related to the fact that students were trying to transfer comprehension strategies from text to video. (Caspi, Gorsky, & Privman, 2005).

The transience of animation and video may aid the development of mental models for dynamic phenomenon (such as physics) (Kozma, 1991). However, individual differences in viewing strategies may have a large effect on knowledge acquisition. Zahn, Barquero & Schwan (2004) observed individual differences in strategies for viewing videos, and found that these strategies were significantly and positively correlated with knowledge acquisition, whereas several specific design characteristics they were investigating did not correlate with knowledge acquisition. Understanding how viewers process animation lags behind its use. There is a lack of understanding of when animations are effective and how best to design them (Reimann, 2003).

Until further research reveals more effective development and utilization strategies for video and animation, its role in assessment should mirror its use in the classroom. The observed differences in learning approaches and comprehension suggest that using video and animation to present novel material could introduce CIV.

For animations see Video/Animation component guidelines and Marcromedia Flash accessibility guidelines: **http://www.macromedia.com/resources/accessibility/flash8/**

If text accompanies video, ensure that the timing and placement of text is commensurate with the reading level of students. Audio should accompany the text whenever possible. (this came out in the student interviews—tchrs and students liked having the text that accompanied the rock formation items, but the reading level was too high, the timing of the text was too fast, and it didn't have audio accompaniment)

> The Video/Animation Component Checklist on page 167 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Video/Animation Component Considerations – Overview	
Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination
Cognitive Processing:	· Visual Processing Skills · Using Graphic Conventions · Using Iconic Conventions · Visual Syntax Fluency · Background Knowledge · Comprehension Strategies · Planning and Organizing Skills · Concentration and Attention
Motoric Processing:	· Object Manipulation and Navigation Abilities
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. Visual Ability

1.1. Audio description is narration accompanying a video or animation that describes the visual elements, action, scene changes, graphics and on screen text. It may be incorporated into multi-media as open (always visible to all users) or closed (visible only to users who turn on the audio description). Synchronized Multimedia Integration Language (SMIL) and Synchronized Accessible Media Interchange (SAMI) formats support the inclusion of audio description. SMIL is played by Quick Time Player, Real Player, and Oratrix GRiNS Player. SAMI is played by Windows Media Player. Flash animation technology allows native captions to be created within an animation. MAGpie (Media Access Generator) can be used to create audio for SMIL. More information about SMIL, SAMI and MAGpie, including techniques and code can be found at http://ncam.wgbh.org/cdrom/guideline/guideline2.html and http://ncam.wgbh.org/richmedia/ .

2. Visual Acuity

2.1. Flash uses vector-based graphics, which allows it to be re-sized without a reduction in clarity. Zoom is not ideal for bitmap-based video or animation as it introduces pixelation, but should be made available.

3. Visual Discrimination

3.1. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006). Students should be able to view animations and videos in monochrome.

Cognitive Processing

Considerations for reducing barriers due to differences in:

4. Visual Processing Skills

4.1. Highlight critical features in videos or animation; provide instructions to direct attention to relevant areas or features of the presentation; reduce animation or video to essential elements.

4.2. Enable pausing, reviewing, and fast forwarding so students can pace their own visual processing.

5. Using Graphic Conventions

5.1. Do not assume knowledge of graphic conventions, provide descriptions and literal depictions.

5.1.1. Provide alt-text and/or explicit directions for graphic conventions used for the video/animation player.

5.1.2. Avoid graphic conventions within videos/animation. If they exist, provide a legend or definitions in advance of the video/animation so students have the relevant knowledge while watching.

6. Using Iconic Conventions

6.1. Avoid using icons as the sole means of conveying content.

6.1.1. Provide a legend for icons encountered during a video/animation. For example, an image of a snowman in an animation is an icon of winter that might not be familiar to students from Latin America.

6.1.2. Use rollovers or mouseovers to provide definitions of icons (rollovers/mouseovers are not inherently accessible to screen readers unless they can be locked and a cursor directed to them).

7. Visual Syntax Fluency

7.1. Reminders of comprehension strategies can be incorporated into test instructions, attached to item directions, or embedded in videos/animation. Research on embedded comprehension and meta-cognitive questions (Tobias, 1987, 1988; Dalton, Pisha, Eagleton, Coyne, & Deysher, 2002; Salomon, Globerson, & Guterman, 1989), and review of comprehension strategies (Anderson, Horney, & Blair, 1999-2001) has been shown to positively impact comprehension of hypertext. There is potential for the same supports to work in video/animation comprehension.

7.2. Providing students with bookmarking as an analog to highlighting of static text and images will provide them opportunities to mark salient features of a video or animation.

8. **Background Knowledge**

   8.1. Links to relevant prior knowledge could be provided alongside a video/animation. Topic knowledge has been shown to impact what and how much is comprehended (Stahl, 1991, 1989; Hirsch, 2003). The extent to which links to material provided during testing, material which is succinct enough to be used within a testing session, can improve performance of students who lacked relevant background knowledge is unknown. Judicious use of such supports is recommended as there is also the potential for confusing or slowing students down, negatively impacting their performance.

9. **Comprehension Strategies**

   9.1. Reminders of viewing and interpretation strategies can be incorporated into test instructions, attached to item directions, or programmed to appear at timed intervals throughout the item interaction.

   9.2. Navigation controls (pause, fast forward, reverse, replay, and search) permit students to determine their own pace and revisit information on an as-needed basis.

10. **Planning and Organizing Skills**

    10.1. Advance organizers, concept maps, and other graphic organizers can support categorization and conceptual understanding, and provide information about structure. They can be presented in advance of content to help activate and organize background knowledge, as a means to organize thoughts and information during viewing, or as post-organizers after students interact with the media. In each instance, the graphic organizer could appear in the item layout, or be made available through a link.

       10.1.1. When presented in advance a graphic organizer introduces the structure, prepares viewers for the content and assists in comprehension.

       10.1.2. During viewing, a graphic organizer can be used as a note taking template, simultaneously reasserting the structure and relationships in the image. Whether the organizer is entirely blank, or contains some level of detail needs to be determined by item developers with respect to construct modification.

10.1.3. Graphic organizers used after viewing videos and animations can be used to summarize and clarify content and structure. They can also be presented blank, or nearly blank, and used by viewers to summarize and clarify their comprehension.

11. Concentration and Attention

11.1. Prompting could be embedded throughout a test, or within particular items by having attentional reminders appear on screen or as an auditory alert according to predetermined time intervals. Brooks, Todd, Tofflemoyer, and Horne (2003) have shown increases in performance due to prompting.

11.2. Simplified content can make important relationships more clear by removing "background noise."

Motoric Processing

Considerations for reducing barriers due to differences in:

12. Object Manipulation and Navigation Abilities

12.1. Provide keyboard alternatives for all on screen navigation and actions. This includes "jog wheel"-like functionality that allows students to control the rate and direction of playback.

12.2. Respect operating system defaults to ensure assistive device and software compatibility.

Executive Processing

Considerations for reducing barriers due to differences in:

13. Goal Setting Ability

13.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

13.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

14. Goal Maintenance and Adjustment

14.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

14.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

14.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

15. Progress Monitoring

15.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

16. Working Memory

16.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

17. Self-regulation

17.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

18. Intrinsic Task-specific Motivation

18.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

18.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

19. Extrinsic Incentives

19.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

19.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

20. Test Conditions

20.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

20.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

20.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

20.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

20.5. Items should be reviewed for age appropriate content.

Response Options

In the context of the UD-CBT Guidelines, response options focus on the processes involved in students' understanding of and responding to test items. Even for traditional "simple" item types such as multiple choice (MC), these are not simple processes. Students in the early elementary grades, and those unfamiliar with the MC format, must be taught how to mark their answers on a bubble sheet, and particularly how to make sure the item number on the bubble sheet corresponds to the item they are answering in the test booklet. Additionally instruction is given about filling in the bubbles fully and precisely. Students of any age who have poor organizational skills, concentration difficulty, or are physically impaired face greater challenge on MC tests than students without these issues.

In the realm of CBT both understanding and responding have the potential to be even more complex than in traditional tests. For example, a simulation environment such as a biology lab in which students dissect a frog, identify organs, and write about the relationship between various systems, presents multiple stimulus and response elements to understand and physical challenges including dexterity, navigation, and keyboarding. Such a virtual laboratory might be a more authentic assessment item than a series of MC items about the function and appearance of frog organs, but it also tests skills and abilities in addition to the knowledge addressed by the MC items.

Item developers need to be aware of the impact of interactive items on student performance. There is an assumption underlying traditional testing that all students who do not have physical disabilities will be equally skilled and, therefore, equally affected by the paper-and-pencil medium. This is not necessarily true, but equal skill and impact is even less tenable in CBT. All students will be challenged when presented with novel item formats. Those with more computer experience, adaptability, and/or the ability to learn an interface quickly will likely have less difficulty answering certain items, incur less anxiety comprehending the interfaces, and will spend less time on interactive items; all of which could bias their scores upwards, or bias the scores of students with less experience and/or ability downwards.

Students with disabilities who use assistive technology (AT) devices present a particular puzzle for interactive CBT items. Due to the nature of their disability, some of these students may be more familiar with working in simulated environments than their peers, and be facile using their device(s) to interact with a computer. However, due to physical disabilities, for example, students may have poor head and eye movement control that may make concentrating on a computer screen difficult, particularly with complex layouts or significant amounts of text. Because it may take extra time for these students to produce controlled movements, they may need extra time on items, and depending on the effort required, the physical demands may impact their cognitive response. For instance, a student for whom head control and eye movement are difficult could become fatigued by a long text passage, which could cause him to misunderstand a MC distractor and answer incorrectly. Students with fine motor control disabilities will be affected by items that require dexterity. The W3C guidelines for Web Accessibility and User Agent accessibility, the NCAM (National Center for Accessible Media) Guidelines, and IMS's guidelines for accessible learning applications, all suggest providing keyboard commands for all mouse actions. Following this suggestion should remove the impact of dexterity on student scores, provided students are aware of the option and know how to use it. Of course, if dexterity is the intended construct of an item, such commands would modify construct measurement.

Deciding which response options guidelines to implement will require input from multiple experts. Many of the tools (such as keyboard commands, explicit labeling, and highlighting links and moveable objects) will support students with disabilities, as well as those without. However, there is a tradeoff between including tools and highlighting elements on the screen. There is always potential for students to become confused or overwhelmed by the choices of tools or complexity of layout that have been included, ostensibly, for their benefit. This must be taken into consideration during test design.

The Response Options Component Checklist on page 168 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Response Options Component Considerations – Overview

Perceptual Processing:	· Visual Acuity · Visual Discrimination · Identifying Stimulus and Response Components · Distinguishing Response Actions
Cognitive Processing:	· Understanding Response Requirements
Motoric Processing:	· Dexterity · Navigation Abilities · Object Manipulation Abilities
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

21. Visual Acuity

21.1. All fonts should allow examinees to adjust size and fonts. If choice is not an option, size should be relative and not fixed.

21.2. For images and icons used to represent tools, or response elements, see Image Components guidelines above.

22. Visual Discrimination

22.1. Black or dark content on white or pastel background generally has the highest readability for most students. For low-vision students, reverse contrast should be made available either through operating system-level features or directly in the test administration software.

22.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

23. Identifying Stimulus and Response Components

23.1. Tools and item elements that appear throughout a test should be familiar to students and used consistently. Use common and familiar icons and labels for them. Provide a tutorial on common tools and the functionality of common item elements immediately in advance of testing, in addition to practice tests available well in advance of testing.

23.2. Explicitly label all tools and item elements using alt-text or image titles. These should be available to all students.

23.3. Provide physical separation (framing) between moveable objects and the areas to which they can be moved, when appropriate.

23.4. Use physical separation (framing) to visually distinguish between stimulus and response elements, and to provide easy navigation by screen readers and single switch devices.

24. Distinguishing Response Actions

24.1. Provide simple, clear instructions using common language.

24.2. Highlight all enabled elements on a page using the common methodology established for the entire test. For example, use the same radio buttons to identify and allow students to select multiple choice responses across all multiple choice items, or use a single text box design for composition throughout a test.

24.3. Provide simultaneous highlighting of elements whose functions are related. For instance, if a student is to drag blocks to a particular area on the screen to build objects, the area on the screen could be highlighted at the same time as the student selects the block to be dragged.

24.4. Use color and/or highlighting in a consistent manner to tie elements together within a layout. For instance, if a magnifying glass is used to click on an image, and the enlarged portion of the image is presented in a separate frame, the enlarged section and the magnifying glass should be tied together by color and/or highlighting. Explicit instructions should convey the relationship as well. Alt-text should convey the ability and resulting action of any tool, such as the magnifying glass.

24.5. There should be a clear and logical relationship between any mouse active behaviors and the corresponding keyboard commands, or assistive device commands. For example, increasing the temperature of matter in a cylinder above a Bunsen burner could accomplished by using the '+' (plus), '>' (greater than), and/or up arrow keys, which could also be accessed by a single switch device.

Cognitive Processing

Considerations for reducing barriers due to differences in:

25. Understanding Response Requirements

25.1. Provide explicit instructions that indicate all steps necessary to finish the item.

25.2. Use tools, objects, and layouts consistently throughout a test.

25.3. Provide practice tests and, potentially, instructional materials that contain environments, objects, and tools that are used in a manner consistent with the assessment. If possible, mimic environments, objects, and tools from instructional materials used throughout the state or testing area.

25.4. For responses that require actions, provide animations that indicate how actions can be completed. Animation should not indicate how to achieve the correct response.

Motoric Processing

Considerations for reducing barriers due to differences in:

26. Dexterity

26.1. Unless dexterity and precision of movement are intended constructs (which would rarely, if ever, be the case), snap-to constraints should be used for all drag and drop activities.

27. Navigation Abilities

27.1. Provide keyboard commands (mouse keys, etc.) for all navigation and mouse actions.

27.2. Ensure tabbed navigation makes sense, which is how Braille devices, screen readers, and many assistive devices navigate a screen.

27.3. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item.

28. Object Manipulation Abilities

28.1. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item. Single switch devices will often use dwell time selection. Consider the impact of dwell time selection on animated screen elements.

28.2. Provide keyboard options to select objects on the screen, select and hold while moving them, and options to release them.

28.3. Provide structured navigation with tabbing to simplify moving through the screen, or completing actions for screen readers, Braille devices, and single switch devices.

Executive Processing

Considerations for reducing barriers due to differences in:

29. Goal Setting Ability

29.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

29.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

30. Goal Maintenance and Adjustment

30.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

30.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

30.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

31. Progress Monitoring

31.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

32. Working Memory

32.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

33. Self-regulation

33.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

34. Intrinsic Task-Specific Motivation

34.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

34.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

35. Extrinsic Incentives

35.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

35.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

36. Test Conditions

36.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

36.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

36.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

36.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

36.5. Items should be reviewed for age appropriate content.

Active Objects and Links

Active objects and links are features unique to CBT that allow examinees to navigate within and between items, to link to tools, and to manipulate objects on screen. The majority of links can be made accessible by providing alt-text that identifies the functionality or destination of a static link, and ensuring the test system is compatible with assistive devices. Many navigational features can be accessed through keyboard commands, or can have customized keyboard commands created for them. The primary challenge for item development with active objects and links is with the objects.

Interactive items that require students to manipulate items or objects on the computer screen present potential challenges for all students, but can be particularly difficult for students with physical or sensory disabilities that impair their mobility or vision. For students who are blind or have very low-vision, it may not be possible to precisely move objects from one location to another on screen. Unless tabbed navigation or keyboard commands to manipulate objects are intuitive, it may be necessary to create alternate equivalent items for these students. Similarly, for students with physical disabilities, precisely manipulating onscreen objects may be difficult, fatiguing or impossible without structured or tabbed navigation (for more about structured navigation see http://www.w3.org/TR/WAI-USERAGENT/guidelines.html#gl-navigation guideline 9). Unless dexterity and precision are the intended measurement constructs (which would rarely, if ever, be the case), snap-to constraints and keyboard commands should be available for all students.

> The Active Objects and Links Component Checklist on page 169 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Active Objects and Links Component Considerations – Overview

Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination · Color Perception · Shape Recognition
Cognitive Processing:	· Visual Processing · Using Active Object Conventions · Hypertext Syntax Fluency · Comprehension Strategies · Concentration and Attention
Motoric Processing:	· Object Manipulation and Navigation Abilities
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. **Visual Ability**

   1.1. All objects, buttons or links should be accompanied by alt-text and long descriptions with d-links. Examples:

   - http://www.w3.org/TR/WCAG10-CORE-TECHS/#text-equivalent
   - http://ncam.wgbh.org/cdrom/guideline/guideline4.html
   - http://www.ncddr.org/du/researchexchange/v02n01/design.html

2. **Visual Acuity**

   2.1. If objects need to be moved on screen, the impact of zoom on the orientation of the objects and whether they are visible on the screen simultaneously should be considered.

2.2. Image and object size should be flexible.

3. **Visual Discrimination**

3.1. Black or dark content on white or pastel background generally has the highest readability for most students. For low-vision students, reverse contrast should be made available either through operating system-level features or directly in the test administration software.

3.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

4. **Color Perception**

4.1. Avoid using hue differences as the sole means of conveying information. Hue differences combined with luminance and/or texture differences are fine.

4.2. Avoid common color blindness combinations; provide monochromatic option, user specified color combinations.

5. **Shape Recognition**

5.1. Avoid using shape as the sole means of conveying information, i.e. shapes as icons.

5.2. Provide alt-text and long descriptions with d-links.

Cognitive Processing

Considerations for reducing barriers due to differences in:

6. **Visual Processing**

6.1. Use client side image maps with alt-text for all links.

6.2. Provide a list of all hypertext links

7. **Using Active Object Conventions**

7.1. Avoid using icons as the sole means of conveying content or directions.

7.1.1. Provide a legend for icons used throughout a test, accessible at any time through a link or toggle page.

7.1.2. Provide customizable icons, i.e. let examinees choose icons to represent tools or commands at the beginning of the test.

7.1.3. Use rollovers/mouseovers to provide definitions of icons (rollovers/mouseovers are not inherently accessible to screen readers unless they can be locked and a cursor directed to them).

8. Hypertext Syntax Fluency

8.1. Use alt-text to describe all links

8.2. Clearly indicate the target of all links

8.3. Allow actions to be reversed and enable reverse navigation.

8.3.1. Preserve prior screen information when students navigate backwards.

8.3.2. Allow students to clear select portions of their work, or last action, without resetting the entire item.

9. Comprehension Strategies

9.1. Provide explicit instructions about the process of interacting with active objects

9.1.1. Model actions and use of tools

9.1.2. Highlight the relationship between action being undertaken and its stage in the overall item requirements.

9.1.3. Use tools and actions consistently throughout the test.

9.1.4. Highlight critical relationships (use color, shadowing, spotlighting, etc.) such as the object to be moved and its target, or provide explicit text that describes the relationship between elements on the screen.

9.1.5. Keep alerts or pop-ups on screen until dismissed so timing does not affect access.

10. Concentration and Attention

10.1. Allow users to access multiple sources of information separately when they are presented simultaneously.

10.2. Provide a focus indicator to identify the elements on which an examinee is acting, the stage of the item, and the active region of the screen.

Motoric Processing

Considerations for reducing barriers due to differences in:

11. Object Manipulation and Navigation Abilities

11.1. Provide keyboard alternatives for all on screen navigation and actions.

11.2. Respect operating system defaults to ensure assistive device and software compatibility.

Executive Processing

Considerations for reducing barriers due to differences in:

12. Goal Setting Ability

12.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

12.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

13. Goal Maintenance and Adjustment

13.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

13.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

13.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

14. Progress Monitoring

14.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

15. Working Memory

15.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

16. Self-regulation

16.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

17. Intrinsic Task-specific Motivation

17.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

17.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

18. Extrinsic Incentives

18.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

18.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

19. Test Conditions

19.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

19.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

19.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

19.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

19.5. Items should be reviewed for age appropriate content.

Constructed Response: Text Composition

Just as text is a ubiquitous form of communication that is predominantin educational materials, composition is commonly used for constructed responses. Its use in non-language arts subject areas is generally not intended to measure general writing ability, language knowledge and use, grammar, spelling, or handwriting. Consider a science item that asks students to write about the difference between sexual and asexual cellular reproduction. For some students conveying that they know the difference is complicated by language or grammar, whereas they might be able to draw or otherwise represent the difference quite easily. In standardized testing, it is not currently possible to let students choose between writing and drawing an answer. However, it is possible to provide supports that help students convey their knowledge without compromising measurement.

For some students, composing text on the computer presents challenges unrelated to their writing ability. For students without disabilities, the research suggests that students who are experienced using computers for writing tend to perform better when writing on a computer than writing long hand (Russell & Haney, 1997; Russell & Plati, 2001; Goldberg, Russell & Cook, 2003). However, a 1999 study in Oregon indicated 7th grade SWD performed less well on a computer composition than on hand-written work, while students without disabilities performed equally well on computer and paper (Hollenbeck, et al, 1999).

There has been little research on the impact of assistive tools on the writing process. Cahalan-Latusis (2004) refers to one study in which some students were allowed to use spell checker and others were not when composing on a word processor. Students in the group allowed to use spell-check received significantly higher scores on organization, sentence fluency, and conventions. No significant differences were found between the two groups on ideas and content, voice, and word choice scores. Dolan et al. (2006) in a pilot study with high school students found improvement in some aspects of compositional writing when students had access to text-to-speech for reading back what they had written.

A study conducted by inTASC (inTASC.org) with funding from the New England Compact Enhanced Assessment Grant more fully investigated the impact of allowing students to use computers for writing and allowing them use of editing tools (Russell, Higgins, & Hoffman, 2004). This study found increased scores on topic development, English standards, and total score for students using computers versus paper, and further score increases in all categories for students using computers with editing tools. The authors investigated the impact of a number of factors on the scores including race, language, school computer use, home computer use, keyboarding, and computer fluency. These factors did affect the score differences, but regardless of the control variables students performed better on computers than on paper, and best when using a computer with editing tools.

Typing-versus-handwriting bias may have confounded the results of a study on the writing ability of college students with disabilities (Higgins & Raskind, 1995). The study found that using a speech recognition system provided a significant benefit compared to writing with no assistance, either by hand or on a word processor. Dictation to a scribe was also compared and no significant difference was found between the scribe condition and speech recognition or no assistance. These researchers conducted another study (Higgins & Raskind, 1995) that investigated the impact of text-to-speech with synchronous highlighting, a human reader, or no assistance on the ability of students with disabilities to proofread and locate errors in an essay they had written previously. The text-to-speech technology resulted in students finding significantly more errors (36%) than the human reader (32%) or the no assistance condition (25%).

The Constructed Response: Text Composition Component Checklist on page 170 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Constructed Response: Text Composition
Component Considerations – Overview

Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination
Linguistic Processing:	· English Language Proficiency · Vocabulary Knowledge · Syntactic Skills
Cognitive Processing:	· Medium Familiarity and Expertise · Planning and Organizing the Writing Task · Writing Fluency
Motoric Processing:	· Production Dexterity · Navigation Abilities · Strength and Mobility · Automaticity
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. Visual Ability

1.1. Providing access to composition for blind students requires multiple options. Some students use Braille input devices, while others use traditional or modified keyboards. Additionally, they may use Braille displays to review their work, screen readers, or a combination. Deaf blind students will only be able to use Braille devices. Providing instruction and practice in how applications or devices will interact with the test system is necessary prior to actual testing.

1.1.1. Screen Readers and Self-Voicing Output: Most screen reading software uses Microsoft Active Accessibility (MSAA) protocols (Microsoft Corporation, 2006); these can be accessed at http://msdn.microsoft.com/library/default.asp?url=/library/en-us/msaa/msaastart_9w2t.asp. Screen readers generally handle text in single columns best, and often have difficulty when multiple hypertext markup language (HTML) frames are used.

1.1.2. Applications which provide internal text-to-speech rendering of text-based files are known as self-voicing text-readers. Capabilities such as reviewing text, reviewing individual words or sections, and checking the spelling of words that are mispronounced or homonyms need to be included in self-voicing applications. Including text-to-speech capabilities in a test application may remove the need to provide assistive technology during testing, but the developer must ensure functionality that matches common screen reading applications and allow test takers to practice using the self-voicing technology.

1.2. Refreshable Braille displays generally present only one line of text at a time, and those with less than 80 cells provide only part of a line at a time. Most portable displays have 40 cells, but several 85 cell displays are available. Some displays have keys that can be programmed for additional functionality, such as mouse key commands. Many are designed to be used in conjunction with screen reading software. Most Braille displays are designed to work with the Windows operating system, applications should not disable operating system (OS) functions to ensure compatibility.

1.2.1. Ensure that screen readers and Braille displays will only present information in the composition areas when students wish to review their composition and do not re-present the entire screen. Use focus indicators when applicable.

2. Visual Acuity

2.1. All fonts used should allow examinees to adjust size and/or be amenable to the use of cascading style sheets (CSS).

2.1.1. As a default, 12 point fonts are considered standard for paper, fonts between 12 and 18 are considered enlarged, and 18 point fonts are considered large print (Allman, 2004).

2.1.2. The impact of font size adjustments on the item's layout should be considered, and any design changes necessary to maintain the general look and feel of the item should be automatically invoked when corresponding font sizes are chosen. Text should be allowed to reflow or rewrap when the font size changes.

3. Visual Discrimination

3.1. Black or dark content on white or pastel background generally has the highest readability for most students. For low-vision students, reverse contrast should be made available either through operating system-level features or directly in the test administration software.

3.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

3.3. Sans-serif fonts (e.g. Verdana, Arial) should be used as they generally have higher readability on-screen than serif fonts (e.g. Times).

Linguistic Processing

Considerations for reducing barriers due to differences in:

4. English Language Proficiency

4.1. Composition in languages other than English should be offered for ESL students, and Braille for blind and deaf/blind students.

5. Vocabulary Knowledge

5.1. Providing vocabulary support with dictionary and/or thesaurus tools. Dictionaries and thesauruses should be available in multiple languages and provide translation to English (or the expected language of composition).

5.1.1. Providing access to an online dictionary through a tool link requires students to exhibit initiative to find definitions. Differences in student motivation and initiative could impact performance and introduce CIV. However, providing a dictionary instead of rollovers/mouseovers allows students to find definitions for words other than those specified by the test developers, which can alleviate CIV due to differences in vocabulary knowledge. Providing a spell checker with the dictionary tool will help students find words that they might spell incorrectly; such a tool is currently built into most web search engines. This would be especially important for looking up words that are heard through audio rather than read.

5.1.2. Deaf students may need to sign a word that they want translated from ASL to English. A human aide/translator would be necessary for this purpose and should be validated as an accommodation if translation to other languages is being provided.

5.1.3. Use appropriate markup language to facilitate pronunciation for screen readers or self-voicing reference tools. See W3C guidelines for examples and techniques (http://www.w3.org/TR/WAI-WEBCONTENT/#Guidelines).

6. Syntactic Skills

6.1. Syntactic skills in composition refer to students' abilities to organize words, use textual elements, and compose in an appropriate format.

6.1.1. Grammar aids can assist with sentence structure.

6.1.2. Graphic organizers or outlines can assist in organization and structure.

6.1.3. Providing self-voicing text-to-speech allows students to hear their writing read aloud, as mentioned above. This can assist students with content and syntax.

Cognitive Processing

Considerations for reducing barriers due to differences in:

7. Medium Familiarity and Expertise

7.1. Ideally students would be able to choose the response option for conveying their knowledge with which they felt most comfortable if writing was not the focus of measurement. However, many students are not able to accurately determine this. Alternately, providing models or templates for formats (letters, lists, etc.), and models of possible responses to similar items, could assist students with the knowledge but less writing ability or experience.

8. Planning and Organizing the Writing Task

8.1. Students should be provided access to the rubric that will be used to judge their composition, even if the composition is a short answer. Note that not all students will choose to refer to the rubric, which could be considered CIV. In addition to the actual rubric, reminders of what the answer should contain can be worked into the prompt. For example: "The picture shows two children on a playground. An angry-looking adult is approaching the children. Imagine and describe what is happening on the playground and what happens next. Be sure to include all the people in the picture in you description. You should write your description in complete sentences, but you will not be scored on spelling or punctuation."

8.2. Graphic organizers are often used to teach writing. They can help students conceptualize and organize their thoughts before beginning to write, and can provide a template for the written response. Note that if a graphic organizer is offered through an optional link, not all students will choose to use it, which could introduce CIV. However, a reminder about planning can be included in the writing prompt. Adding to the previous example "…You may want to create an outline of what you plan to write and organize what think is going to happen in the order it occurs."

9. Writing Fluency

9.1. Students overly concerned about spelling, grammar, and punctuation may be interrupted in their writing process and impeded in their ability to convey content. Providing spelling and grammar tools can alleviate worry and let students concentrate on conveying their ideas in situations where such skills are construct-irrelevant. Additionally, reminders about how to write fluently could be included in the prompt, presented by an avatar or virtual mentor, or provided via a link to writing strategy support.

Motoric Processing

Considerations for reducing barriers due to differences in:

10. Production Dexterity

10.1. There are several types of physical disabilities that impact the ability to type. Students with fine motor control issues will likely compose using a keyboard, but they work more slowly than students without motor challenges. These students may become fatigued from the additional physical effort typing requires, which can affect their writing. They may benefit from word prediction. Students with dysgraphia, dyspraxia, moderate cognitive disabilities, autism spectrum disorders, and attention deficit/hyperactivity disorder, may have comorbid fine motor challenges.

10.2. Students with gross motor disabilities are more likely to use assistive devices to interface with the computer. Composition is often fatiguing for students with gross motor issues both because of the physical impact of their disability, and the tedium of typing with assistive devices. Word prediction can also benefit students with gross motor disabilities.

10.3. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item

11. Navigation Abilities

11.1. Keyboard commands should be made available for all mouse behaviors.

11.2. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item

11.3. Allow students to choose between scrolling and paging through their composition.

12. Strength and Mobility

12.1. Provide flexibility in the physical setting. Many students with physical disabilities may not be comfortable in a traditional CBT setting. They may also need changes in posture, or to be able to rest from maintaining the posture necessary for interacting with the computer.

12.2. Elimination of time limits and/or providing timing accommodations can help reduce barriers for students who might otherwise experience fatigue due to strength and mobility challenges.

13. Automaticity

13.1. Students have varying levels of comfort and practice with writing. Students who are not automatic writers may prefer to be able to draft ideas, outlines or versions of their composition. Reducing or eliminating time limits and/or providing timing accommodations allows students who need more time for planning, writing and reviewing are not prevented from it, or threatened by the pressure of a time limit.

Executive Processing

Considerations for reducing barriers due to differences in:

14. Goal Setting Ability

14.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

14.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

15. Goal Maintenance and Adjustment

15.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

15.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

15.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

16. Progress Monitoring

16.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

17. Working Memory

17.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

18. Self-regulation

18.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

19. Intrinsic Task-specific Motivation

19.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

19.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

20. Extrinsic Incentives

20.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items “share” in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

20.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

21. Test Conditions

21.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students’ IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

21.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

21.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

21.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

21.5. Items should be reviewed for age appropriate content.

Constructed Response: Math

Constructed response math items face many implementation challenges. Mathematical symbols and functions do not appear on standard keyboards, which requires an on-screen math keyboard or stylesheet. These tools may be confusing or physically challenging for many students to use. An on-screen calculator that generates mathematical expressions is the most viable option for most constructed responses, but it does not work responses with math and text interspersed. It is also insufficient for some complex mathematics, and for very simple mathematics. Although the National Council of Teachers of Mathematics (NCTM) has advocated that calculators be fully integrated into assessments at all grades (1998), many states still want to measure basic calculation. However, it is difficult to represent some basic calculation on screen without providing clues to the answer or process; consider long division as an example.

Composing math constructed response items using assistive devices adds another level of challenge. For students who use screen readers or Braille devices to be able to review the work they have entered, it must be encoded in MathML or LaTeX, and even this does not ensure accessibility. MathML is not yet supported by many browsers, but, according to NCAM, it is useable by many assistive technologies. LaTeX can be converted to Nemeth Code for Braille displays, and there is a project in development that will allow students to enter Nemeth Code and display math in LaTex. However, these existing and in-development LaTeX (http://www.maths.tcd.ie/~dwilkins/LaTeXPrimer/)and Nemeth technologies are designed to convert whole documents and not for interactive environments. MathML is a World Wide Web Consortium (W3C) specification based on XML, and more information is available at http://www.w3.org/TR/MathML2/.

The Constructed Response: Math Component Checklist on page 171 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Constructed Response: Math Component Considerations – Overview	
Perceptual Processing:	· Visual Ability · Visual Acuity · Visual Discrimination
Linguistic Processing:	· Creating Graphs or Tables · Creating Diagrams or Drawings · Creating Equations
Cognitive Processing:	· Calculations · Problem Solving
Motoric Processing:	· Production Dexterity · Navigation Abilities · Strength and Mobility · Automaticity
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. Visual Ability

1.1. Providing access to math for blind students is particularly complex for constructed response items. Some students use Braille input devices, while others use traditional or modified keyboards. Braille devices can input Nemeth code, but LaTeX (http://www.maths.tcd.ie/~dwilkins/LaTeXPrimer/) coding is required for the entered text to be accessible to a screen reader or Braille output device. Providing instruction and practice in how applications or devices will interact with the test system is necessary prior to actual testing.

1.1.1. Screen Readers and Self-Voicing Output: Most screen reading software uses Microsoft Active Accessibility (MSAA) protocols (Microsoft Corporation, 2006); these can be accessed at http://msdn.microsoft.com/library/default.asp?url=/library/en-us/msaa/msaastart_9w2t.asp. Screen readers generally handle text in single columns best, and often have difficulty when multiple hypertext markup language (HTML) frames are used.

1.1.2. Applications which provide internal text-to-speech rendering of text-based files are known as self-voicing text-readers. MathML or LaTeX coding would also be necessary for an internal reader.

1.1.3. Ensure that screen readers and Braille displays will only present information in the composition areas when students wish to review their composition and do not re-present the entire screen. Use focus indicators when applicable.

1.1.4. AsTeR is an application that reads Latex notation, creates audio of the mathematical expression, and allows the audio to be navigated. (http://www.cs.cornell.edu/Info/People/raman/aster/aster-toplevel.html)

1.1.5. HP EzMath is a notation for embedding mathematical expressions in web pages, based on how expressions are read aloud. It will soon be available as open source at www.sourceforge.net.

1.1.6. IBM TechExplorer (http://www.software.ibm.com/techexplorer/) is a plug-in for Navigator and Internet Explorer, an IE 5.5 XML Behavior, and an ActiveX control for applications like Microsoft PowerPoint and Word. TechExplorer enables the display of Tex Latex and MathML documents and the publishing of interactive scientific material on the Web. Version 3.1 includes full support for MathML 2.0, augmented Latex display, options for scripting/programming and a web equation editor. It has a Macintosh version, Mathematica connectivity, and Internet Explorer MathML Behavior (IMS Design Guidelines: http://www.imsglobal.org/accessibility/accwpv0p6/imsacc_wpv0p6.html#1273818)."

1.1.7. Triangle is an application for blind computer uses which allows them to read, write and manipulate scientific text, do scientific computations, and read graphs and figures. (http://www.december.com/cmc/mag/1998/feb/gartri.html)

1.2. When possible mathematical expressions that are graphical in nature files should be represented using scaleable vector graphics (SVG) or an equivalent, since images can be scaled without degradation. Additional description for screen readers can be provided using appropriate language (NCAM guidelines for spoken mathematics, Mathspeak for Nemeth Code (http://www.rit.edu/~easi/easisem/talkmath.htm). SVG supports text descriptions, stylesheets, DOM2, assistive technologies and input devices.

2. Visual Acuity

2.1. All fonts used should allow examinees to adjust size and/or be amenable to the use of cascading style sheets (CSS).

2.1.1. As a default, 12 point fonts are considered standard for paper, fonts between 12 and 18 are considered enlarged, and 18 point fonts are considered large print (Allman, 2004).

2.1.2. The impact of font size adjustments on the item's layout should be considered, and any design changes necessary to maintain the general look and feel of the item should be automatically invoked when corresponding font sizes are chosen. Text should be allowed to reflow or rewrap when the font size changes.

3. Visual Discrimination

3.1. Black or dark content on white or pastel background generally has the highest readability for most students. For low-vision students, reverse contrast should be made available either through operating system-level features or directly in the test administration software.

3.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

3.3. Sans-serif fonts (e.g. Verdana, Arial) should be used as they generally have higher readability on-screen than serif fonts (e.g. Times).

Linguistic Processing

Considerations for reducing barriers due to differences in:

4. Creating Graphs or Tables

4.1. Generating tables or graphs is best accomplished by providing templates that students adjust to reflect data accurately, or by allowing students to input data into a spreadsheet and having the graph or table automatically generated. Which option is used will depend on whether students are being assessed on accurately representing given data, or solving a problem to generate the data that creates a graph or table.

4.2. Students should be provided with graph paper to draft responses and problem solve. The feasibility of on-screen graph paper for drafting should be tested, with particular consideration for attaching code to student input that makes it accessible to screen readers and assistive devices for students to be able to review and edit (see tables and graphs guidelines).

5. Creating Diagrams or Drawings

5.1. There are good models for drawing palettes in many software applications, for which accessibility has already been considered. Microsoft Word and numerous Adobe products have palettes with drawing tools, plotting tools, line tools, shape palettes, and more.

6. Creating Equations

6.1. Math Type and Equation Editor are software programs with complex palette systems for writing equations that are cumbersome and awkward to use, but they have extensive palettes and functionality for complex mathematics and could be a model for palettes in a math test. Any tool or palette system would require instruction in use and practice before it could be used in an assessment.

6.2. On screen scientific calculators could be used to generate stand-alone equations, and could be linked to a text box so equations could appear in text. Making the text and equations reviewable by screen readers and Braille devices would require complex coding.

6.2.1. The Audio-Accessible Graphing Calculator (http://dots.physics.orst.edu/calculator.html) is a self-voicing Windows application that has been under development and testing for some time by the Science Access Project at Oregon State University. It includes the capabilities to:

- Compute and display visually either of two functions, their sum, or their difference.
- Display the above as an audio tone plot.

- Permit piece-by-piece audio browsing.
- Print the above to any Windows printer including the Tiger tactile graphics embosser.
- Be used as a universally usable on-screen scientific calculator.
- Be used as a powerful expression evaluator.
- Input tabulated data for display.
- Compute statistical functions for tabulated data.

6.2.2. Braille 'n Speak, a Braille calculator that can do standard math and generate tactile graphs (http://www.dinf.org/csun_99/session0113.html). Additional software is available for students needing the power of a financial calculator and graphing calculator functionality is available for the blind using Graphit in conjunction with any Braille embosser.

6.2.3. Triangle is an application for blind computer uses which allows them to read, write and manipulate scientific text, do scientific computations, and read graphs and figures. (http://www.december.com/cmc/mag/1998/feb/gartri.html)

Cognitive Processing

Considerations for reducing barriers due to differences in:

7. Calculations

7.1. Providing on-screen calculators makes the content of items the focus of measurement, not the underlying calculations. If calculation is the construct being measured, access to the calculator should be barred. Some students may prefer stand alone calculators and/or scrap paper.

7.1.1. The Audio-Accessible Graphing Calculator (http://dots.physics.orst.edu/calculator.html) is a self-voicing Windows application that has been under development and testing for some time by the Science Access Project at Oregon State University. It includes the capabilities to:

- Compute and display visually either of two functions, their sum, or their difference.
- Display the above as an audio tone plot.
- Permit piece-by-piece audio browsing.

- Print the above to any Windows printer including the Tiger tactile graphics embosser.
- Be used as a universally usable on-screen scientific calculator.
- Be used as a powerful expression evaluator.
- Input tabulated data for display.
- Compute statistical functions for tabulated data.

7.1.2. Braille 'n Speak, a Braille calculator that can do standard math and generate tactile graphs (http://www.dinf.org/csun_99/session0113.html). Additional software is available for students needing the power of a financial calculator and graphing calculator functionality is available for the blind using Graphit in conjunction with any Braille embosser.

7.1.3. Triangle is an application for blind computer uses which allows them to read, write and manipulate scientific text, do scientific computations, and read graphs and figures. (http://www.december.com/cmc/mag/1998/feb/gartri.html)

7.2. If calculations are irrelevant to the intended construct, the numeric answers could be provided for the student to use in constructing an argument or proof.

8. Problem Solving

8.1. A virtual mentor or avatar could guide students through the steps in the problem, or provide them with a template for their response.

8.2. Provide the argument or proof and require the student to fill in the calculations.

Motoric Processing

Considerations for reducing barriers due to differences in:

9. Production Dexterity

9.1. There are several types of physical disability that impact keyboard access. Students with fine motor control issues will likely compose using a keyboard, but they work more slowly than students without motor challenges. These students may become fatigued from the additional physical effort typing requires, which can affect their problem solving, calculations, concentration, and memory. Students with dysgraphia, dyspraxia, moderate cognitive disabilities, autism spectrum disorders, and attention deficit/hyperactivity disorder, may have comorbid fine motor challenges.

9.2. Students with gross motor disabilities are more likely to use assistive devices to interface with the computer. Complex responses are often fatiguing for students with gross motor issues both because of the physical impact of their disability, and the tedium of typing with assistive devices.

9.3. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item.

10. Navigation Abilities

10.1. Keyboard commands should be made available for all mouse behaviors.

10.2. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item.

10.3. Allow students to choose between scrolling and paging through their composition.

11. Strength and Mobility

11.1. Provide flexibility in the physical setting. Many students with physical disabilities may not be comfortable in a traditional CBT setting. They may also need changes in posture, or to be able to rest from maintaining the posture necessary for interacting with the computer.

11.2. Elimination of time limits and/or providing timing accommodations can help reduce barriers for students who might otherwise experience fatigue due to strength and mobility challenges.

12. Automaticity

12.1. Students have varying levels of comfort and practice with mathematical expression. Reducing or eliminating time limits and/or providing timing accommodations allows students who need more time for planning, solving and reviewing are not prevented from it, or threatened by the pressure of a time limit.

Executive Processing

Considerations for reducing barriers due to differences in:

13. Goal Setting Ability

13.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

13.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

14. Goal Maintenance and Adjustment

14.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

14.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional "reminders" along with the item, or the articulation of longer components into shorter ones with sub-goals or "way-stations" can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

14.3. Some students "perseverate" in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

15. Progress Monitoring

15.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

16. Working Memory

16.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

17. Self-regulation

17.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

18. Intrinsic Task-specific Motivation

18.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

18.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

19. Extrinsic Incentives

19.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

19.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

20. Test Conditions

20.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

20.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

20.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

20.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

20.5. Items should be reviewed for age appropriate content.

Multi-stage/Multi-part Items

Multi-stage items are those defined as requiring students to engage in multiple actions and/or multiple responses; in many cases screen elements or environment change at each stage of multi-stage items. Multi-part items are defined as requiring students to switch between screens to view the entire item. Since the two categories often overlap with the distinction between them often unclear, they will be considered together.

Multi-stage/multi-part items introduce a number of factors that must be considered to ensure their use doesn't introduce construct-irrelevant variance due to their complexity. Furthermore, to the extent that the use of multi-stage/multi-part has to potential to improve the authenticity of an item, it is important that the designs carefully consider the ways in which students are likely to use them.

One of the most critical design concerns is that students are aware of the entire scope of the item and that what they see at any given time is a subset of the item. It is also important that students understand where in a multi-stage item they currently are, and that the impact of returning to an earlier stage and modifying a response is well thought out and communicated to the student.

> The Multi-stage/Multi-part Items Checklist on page 172 provides a quick reference into the sources of construct-irrelevant variance in item designs and the design solutions that can address them.

The following table provides an overview of the sources of variance for the categories of processing relevant to this component:

Multi-stage/Multi-part Items – Overview	
Perceptual Processing:	· Visual Ability · Visual Acuity · Distinguishing Intra-item Navigation Actions · Identifying Stimulus and Response Components
Linguistic Processing:	· Constructing Meaning from Text
Cognitive Processing:	· Understanding Response Requirements · Planning and Organizing Skills · Hypertext Syntax Fluency
Motoric Processing:	· Navigation Abilities
Executive Processing:	· Goal Setting Ability · Goal Maintenance and Adjustment · Progress Monitoring · Working Memory
Affective Processing:	· Self-regulation · Intrinsic Task-specific Motivation · Extrinsic Incentives · Test Conditions

Perceptual Processing

Considerations for reducing barriers due to differences in:

1. Visual Acuity

1.1. All fonts should allow examinees to adjust size and fonts. If choice is not an option, size should be relative and not fixed.

1.2. For images and icons used to represent tools, or response elements, see Image Components guidelines above.

2. Visual Discrimination

2.1. Black or dark content on white or pastel background generally has the highest readability for most students. For low-vision students, reverse contrast should be made available either through operating system-level features or directly in the test administration software.

2.2. Screen contrast should be adjustable and should remain adjustable throughout the test. A free extension for Mozilla Firefox that can be used as a design tool to analyze color contrast between foreground and background can be accessed at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php#comment2 (JuicyStudio, 2006).

3. Distinguishing Intra-item Navigation Actions

3.1. Highlight all enabled intra-item (i.e. between phases and parts) navigation elements using the common methodology established for the entire test. For example, use the same controls for moving to previous and next stage or part, and the same indicator for indicating current position.

3.2. Use color and/or highlighting in a consistent manner to tie stages and part together within a layout.

3.3. There should be a clear and logical relationship between any mouse active behaviors and the corresponding keyboard commands, or assistive device commands.

4. Identifying Stimulus and Response Components

4.1. Ensure stimuli and response components can be distinguished consistently so that students can clearly differentiate the two when navigating between stages and parts.

4.2. Explicitly label all tools and item elements using alt-text or image titles. These should be available to all students.

4.3. Provide physical separation (framing) between moveable objects and the areas to which they can be moved, when appropriate.

4.4. Use physical separation (framing) to visually distinguish between stimulus and response elements, and to provide easy navigation by screen readers and single switch devices.

Linguistic Processing

Considerations for reducing barriers due to differences in:

5. Constructing Meaning from Text

5.1. Support students' understanding of the entire scope of the item by providing clear, concise textual descriptions of how many stages and/or parts exist.

5.2. Further support students' understanding of the scope of the item by providing visual and/or auditory representations of the various stages and/or parts of the item.

Cognitive Processing

Considerations for reducing barriers due to differences in:

6. Understanding Response Requirements

6.1. Provide explicit instructions that indicate all steps necessary to finish the item.

6.2. Use tools, objects, and layouts consistently throughout a test.

6.3. Provide practice tests and, potentially, instructional materials that contain environments, objects, and tools that are used in a manner consistent with the assessment. If possible, mimic environments, objects, and tools from instructional materials used throughout the state or testing area.

7. Planning and Organizing Skills

7.1. Allow actions to be reversed and enable reverse navigation.

7.1.1. Preserve prior stage and part response information for modification when students navigate backwards.

7.1.2. Allow students to clear select portions of their work, or last action, without resetting the entire item.

8. Hypertext Syntax Fluency

8.1. Use alt-text to describe all links to the extent that they are used to provide navigation between item stages or parts.

8.2. Clearly indicate the target of all links.

Motoric Processing

Considerations for reducing barriers due to differences in:

9. Navigation Abilities

9.1. Provide keyboard commands (mouse keys, etc.) for all intra-item (i.e. between stages and parts) navigation and mouse actions.

9.2. Ensure tabbed navigation makes sense, which is how Braille devices, screen readers, and many assistive devices navigate a screen.

9.3. Do not disable any operating system-level functionality. If overriding operating system-level functionality is necessary, provide alternative functionality to designate how the devices' focus indicators should move through the item.

Executive Processing

Considerations for reducing barriers due to differences in:

10. Goal Setting Ability

10.1. Explicit instructions and/or initial prompts can be provided in order to make the purpose or goal for an item more salient or evident. For some students, the fact that the actual purpose or goal for interacting with a component in an item is often implicit rather than explicit is an impediment to processing information strategically – and thus an impediment to comprehension, recall, selectivity, problem-solving (e.g.Schunk & Zimmerman, 1997). Whether such prompts are appropriate for a particular item will depend on whether determining goals (by the student) for reading are construct-relevant.

10.2. A second type of support is more "generic" – not explicitly instructing a student what the goal or purpose for an item component is, but prompting them, and/or scaffolding them, to take the initial step of setting their own purpose or goal for interacting (e.g. comprehending) with an item. Such supports can guide students to focus their efforts and to set goals that are timely and realistic for the type of item in which they will be engaged (Paris & Paris, 2001; Paris, Byrnes & Paris, 2001).

11. Goal Maintenance and Adjustment

11.1. Particularly in complex or extended items, some students will have difficulty in maintaining a consistent goal or purpose for their activity (some, for example, will find themselves continually interrupted by sub-tasks like decoding text, others by irrelevant tasks or distractions in the environment). Where competence in the various sub-tasks are not construct-relevant (e.g. decoding a text, identifying unfamiliar vocabulary), the scaffolding of these subtasks (e.g. providing text-to-speech or links to definitions) is a key to supporting students in maintaining higher-level strategic goals (Dalton et al, Dalton and Rose, 2002).

11.2. Where the ability to sustain goal direction is not construct-relevant, the embedding of optional “reminders” along with the item, or the articulation of longer components into shorter ones with sub-goals or “way-stations” can provide scaffolding to help students sustain goals across longer components and items (Wood, 1988; 2002).

11.3. Some students “perseverate” in striving for goals that are unattainable or inappropriate in the light of relevant feedback (Stone & May, 2002, Dreisbach and Goschke, 2004). Explicit instructions and prompts, especially as feedback following poor initial performance, can make that feedback more explicit and salient, raising the likelihood it will serve as a cue for revising goals and strategies (Butler & Winne, 1995).

12. Progress Monitoring

12.1. In a CBT environment (as opposed to a print environment) it is possible to provide explicit feedback – in alternative and accessible formats - on progress toward goals and in sustaining appropriate effort, making adjustments in ineffective strategies, and terminating effort when goals have been reached. The main purpose of making progress monitoring more explicit is that some students do not effectively monitor their own progress, an executive function, and thus are not able to act strategically, revise plans on the basis of feedback, and so forth. Two broad kinds of progress monitoring scaffolds can be embedded. The first does not involve feedback on performance per se, but locates the student in the overall task structure or text of the item – e.g. a graphic that indicates how many steps have been achieved in a multi-step problem that involves reading. The second kind does involve performance assessment, and provides an ongoing and usually graphic display of progress (three responses right out of four) (Deno, 1999, Palinscar and Brown, 1984).

13. Working Memory

13.1. From a UDL framework there are several different aspects of working memory that should be addressed. When the ability to hold previous information in memory is not construct relevant, provide external memory aids – notebook, checklists, links to explicit information – that can provide support. When the ability to maintain and execute a sequences of actions (e.g. the steps in a recipe or routine) is not construct relevant, provide external organizers or templates (e.g. a timeline, embedded reminders, navigation prompts, sequential templates) (Deshler and Schumaker, many references). When the ability to maintain a goal or incentive is not construct relevant, see below. Such supports can be provided externally to the CBT (e.g. a paper notebook) or embedded within it (an electronic notebook) and implemented as appropriate to individual and construct.

Affective Processing

Considerations for reducing barriers due to differences in:

14. Self-regulation

14.1. General supports for self-regulation – the ability to sustain motivation over extended items or clusters of items, to respond effectively to threats and challenges in items, to manage anxiety – would typically be addressed at the test level rather than at the item level. Within longer items (a long text passage or multi-step item, for instance) it is possible to embed self-regulatory prompts and scaffolds that can guide and support students who are unable to self-regulate or whose emotional state (e.g. test anxiety) limits their effectiveness independently (Lewis, 2004, Goldsmith and Davidson, 2004).

15. Intrinsic Task-specific Motivation

15.1. Where specific content is not construct relevant, alternative contents may reduce content-specific threats to validity that arise from difference in interests and preferences. For example, in assessing comprehension or problem solving, it is advantageous to provide two (or more) different contents or contexts (e.g baseball versus ballet) that differ in superficial features but which maintain equivalent levels of difficulty in construct-relevant features. Such alternatives provide opportunities not only to address background knowledge differences but also the differential effects of motivation, familiarity, and interest.

15.2. The importance of providing alternative representations has been emphasized for perceptual, linguistic, cognitive and executive reasons. Many of those same alternatives provide an additional advantage: addressing the differential emotional reactions to items by students with different histories and abilities in learning and testing environments. Recent cognitive neuroscience research has identified the differential motivational effects of two perceived conditions – threat and challenge. A task, activity, or problem is seen as *challenging* (with accompanying physiological responses) when an individual perceives that they have the cognitive and/or emotional resources they need, even though the task may be difficult. On the other hand, the same task is seen as a *threat (with different physiological responses that prepare for flight)* when the individual perceives that they do not have the mental resources they will need (e.g Blascovich et al, 2000). In the CBT environment it is possible to provide additional resources externally – like TTS for decoding, links for difficult vocabulary, alternative sizes or types of images, etc. – that can reduce a threatening problem to a challenging one, with consequences for achievement and effort (e.g. Fredrickson and Branigan, 2005).

16. Extrinsic Incentives

16.1. Most extrinsic incentives and rewards for performance are delivered within the context of the overall assessment – i.e. not at the item level. To that extent, items "share" in the effects of overall incentives, but such incentives are not typically a part of item construction. Since extrinsic rewards have differential, and often deleterious (Wang and Guthrie, 2004), effects on item comprehension, the uniform use of extrinsic rewards by all students represents a considerable threat to validity. Providing alternatives to any external incentives – positive or negative – is an important consideration in order to reduce the construct irrelevant effects on individual items.

16.2. The *type* of incentives (rewards, social comparison, punishments) used in motivating performance also has differential effects for different types of students as they engage in items and components. While typically administered at the overall assessment level, the differential threats to validity that they confer on individual items should be considered and alternative types of rewards and punishments should be provided or encouraged (Wang and Guthrie, 2004).

17. Test Conditions

17.1. Alternative settings and conditions are common accommodations for paper-based testing and should be incorporated into CBT as well. Decisions about extended time, multiple testing sessions, or alternate locations are part of students' IEPs. If the validity of the test does not depend on time constraints, finishing in one session, or location, these options could be offered to all students without compromising the test., and could improve student performance (Cahalan-Laitusis, 2004; Camara, W.J., Copeland, T., & Rothschild, B., 1998; Ziomek & Andrews, 1998).

17.2. Item bias is CIV that consistently increases or decreases the likelihood that individuals who are members of certain groups will respond correctly to an item. The item and test development process should incorporate multiple means of detecting item bias.

17.3. Involve experts and stakeholders in item reviews to look for racial, cultural, ethnic, and gender bias in item content. Reviews should occur before and field testing and after when response data can be used in the review.

17.4. Items should undergo statistically based Differential Item Functioning (DIF) analyses.

17.5. Items should be reviewed for age appropriate content.

References

All Things Web. (1997). *The Art of ALT*. Retrieved May 5, 2006, from http://www.pantos.org/atw/35534.html

Allan, J. M., Bulla, N., & Goodman, S. A. (2003). *TEST ACCESS: Guidelines for Computer Administered Testing*. Louisville, Kentucky: American Printing House for the Blind.

Allman, C. B. (2004). *Test access: Making tests accessible for students with visual impairments: A guide for test publishers, test developers, and state assessment personnel*. Louisville, KY: American Printing House for the Blind.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bejar, I., Lawless, R., Morley, M., Wagner, M., Bennett, R. & Revuelta, J. (2003). A feasibility study of on-the-fly item generation in adaptive testing. *Journal of Technology, Learning and Assessment*, 2(3). Available from http://www.jtla.org.

Berninger, V. Abbott, R., Jones, J. & Wolf, B. Early Development of Language by Hand: Composing, Reading, Listening, and Speaking Connections. *Developmental Neuropsychology*, 2006, 29. 61-92.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. Forgas (Ed.), *Feeling and thinking: Therole of affect in social cognition* (pp. 59–82). Paris: Cambridge UniversityPress.

Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions*, 5(3), 144-152.

Burling, K.; Beck, R.; Jude, J., Murray, E., Harms, M.; & Dolan, B. (2006). Constructing innovative computer-administered tasks and items according to universal design: illustrative examples with pilot data. Paper presented at the *Annual Meeting of the National Council on Measurement in Education*, San Francisco, CA.

Butler, D.L. & Winne, P.H. Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, Vol. 65, No. 3 (Autumn, 1995) , pp. 245-281

Byrnes, H. (1984). *The role of listening comprehension: A theoretical base*. Foreign Language Annals , 17 , 317-329.

Cahalan-Laitusis, C. (2004). *Accommodations on High-stakes Writing Testsfor students with Disabilities*. (ETS Research Report RR-04-13). Princeton, NJ: Educational Testing Service.

References

California Community Colleges. (1999). *Distance education: Access guidelines for students with disabilities*. Available at: http://www.htctu.net/publications/guidelines/distance_ed/distedguidelines.pdf

Camara, W., Copeland, T. & Rothschild, B. (1998). *Effects of extended time on the SAT I: Reasoning test score growth for students with learning disabilities.* (College Board Rep. No. 98-7). New York: College Entrance Examination Board.

Clapper, A. T., Morse, A. B., Lazarus, S. S., Thompson, S. J., & Thurlow, M. L. (2005). *2003 State policies on assessment participation and accommodations for students with disabilities* (Synthesis Report 56). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Clark, J. (2002). *Building Accessible Websites (serialization*). Available at: http://joeclark.org/book/sashay/serialization/home.html

Corda. (n.d.) PopChart[D]. http://corda.com/d/

Cumming, J. J., & Elkins, J. (1999). Lack of automaticity in the basic addition facts as a characteristic of arithmetic learning problems and instructional needs. *Mathematical Cognition*, 5(2), 149-180.

DAISY Consortium. (2006). *Digital Accessible Information System (DAISY)*. Retrieved May 5, 2006, from http://www.daisy.org

Dalton, B., Pisha, B., Eagleton, M., Coyne, P., & Deysher, S. (2002). *Engaging the Text: Final Report to the U.S. Department of Education*. Peabody, MA: CAST.

Dolan, R.P. Hall, T.E. & Banerjee, M. (2006). *Evaluation of text-to-speech as a technology-based writing support for students with learning disabilities*. In preparation.

Dolan, R. P. (2006). *Evaluating use of graphics by struggling readers through eye tracking*. Paper presented at the AERA Annual Meeting, San Francisco, CA.

Dreisbach, G. & Goschke, T. (2004) How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility, *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Fredrickson, B.L. & Branigan, C. (2005) Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion*, 19, 3, 313-332.

Geary, D. C. (1993). Mathematical disabilities: cognitive, neurospychological, and genetic components. *Psychological Bulletin*, 114(2), 345-362.

Goldsmith, H.H. & Davidson, R. J. (2004) Disambiguating the Components of Emotion Regulation. *Child Development*, 75, 2, 361-365.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, 37(1), 4-15.

Haladyna, T. (1999). *Developing and validating multiple choice items*. Hillsdale, NJ: Erlbaum.

Halydyna, T.M. & Downing, S.M. (2004). *Construct irrelevant variance in high-stakes testing*. Educational Measurement: Issues and Practice 23(1). 17-27.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 13(3), 309-344.

Handel, S. (1993) *Listening: An Introduction to the Perception of Auditory Events*. MIT Press, Cambridge, MA.

Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, 93(3), 615-626.

Harlen, W. & Deakin Crick, R. (2002). A *systematic review of the impact of summative assessment and tests on students' motivation for learning*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Harms, M.; Burling, K.; Hanna, E.; & Dolan, B. (2006). Constructing innovative computer-administered tasks and items according to universal design: establishing guidelines for test developers. Paper presented at the *Annual Meeting of the National Council on Measurement in Education*, San Francisco, CA.

Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *Journal of Technology, Learning, and Assessment*, 3(4). Available from http://www.jtla.org

Hirsch, E. D., Jr. (2003). Reading comprehension requires knowledge-of words and the world. *American Educator*, 27(1), 10-13, 16-22, 28-29, 48.

IMS Global Learning Consortium. (2002). *Guidelines for Developing Accessible Learning Applications.* Retrieved May 5, 2006, from http://www.imsproject.org/accessibility/accv1p0/imsacc_guidev1p0.html

JuicyStudio. (n.d.). *Colour Contrast Analyser Firefox Extension*. Available at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php

JuicyStudio. (n.d.). *Readability statistics analysis*. Available at http://juicystudio.com/article/colour-contrast-analyser-firefox-extension.php

Keenan, J.M., Betjemann, R.S., Wadsworth, S.J., DeFries, J.C. & Olson, R.K. (2006) Genetic and environmental influences on reading and listening comprehension; *Journal of Research in Reading,* 29, 75.

Leighton, J. P. & Gokiert, R.J. (2005). The cognitive effects of test item features: Informing item generation by identifying construct irrelevant variance. Paper presented at the A*nnual Meeting of the National Council on Measurement in Education*. Montreal, Quebec, Canada.

Macromedia (2006). *Flash Accessibility Guidelines*. San Jose, CA: Adobe Systems. Available at http://www.adobe.com/resources/accessibility/flash8/

Messick, S. (1993). Validity. In R. Linn (Ed.), *Educational measurement (3rd edition*, pp.335-366). Phoenix, AX: The Oryx Press.

Messick, S. (1994). *Alternative modes of assessment, uniform standards of validity*. ETS RR-94-60. Princeton, NJ: Educational Testing Service.

Microsoft Corporation. (2006). *Microsoft Active Accessibility*. Retrieved May 5, 2006, from http://msdn.microsoft.com/library/default.asp?url=/library/en-us/msaa/msaastart_9w2t.asp

Millman, J. & Greene, J. (1993).The specification and development of tests of achievement and ability. In R. Linn (Ed.), *Educational measurement (3rd edition*, pp.335-366). Phoenix, AX: The Oryx Press.

Myles, B. S., & Simpson, R. L. (2001). Effective practices for students with Asperger Syndrome. *Focus on Exceptional Children*, 34(3), 1-14.

NARAP. (2006). *National Accessible Reading Assessment Projects*. Retrieved May 5, 2006, from http://www.narap.info/

NCAM. (2003). *Making Educational Software and Web Sites Accessible: Design Guidelines Including Math and Science Solutions*. Retrieved May 5, 2006, from http://ncam.wgbh.org/cdrom/guideline/

NIMAS. (2006). *National Instructional Materials Accessibility Standard*. Retrieved 2006, May 5, from http://nimas.cast.org/

Nunan, D., & Miller, L. (1995). *New ways in teaching listening*. Alexandria, VA: TESOL.

Osaka, N., Osakab, M., Kondoa, H., Morishitaa, M., Fukuyamac, H. & Shibasaki, H. (2004) The neural basis of executive function in working memory: an fMRI study based on individual differences. *NeuroImage,* 21:2, 623-631.

Ostad, S. A. (1998). Developmental differences in solving simple arithmetic word problems and simple number-fact problems: a comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition*, 4(1), 1-20.

Palinscar, A.S. & Brown, A.L. (1984) *Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. Cognition and instruction*, 1984, 1, 2, 117-175.

Pragma-NL. (2002). *Signing Books for the Deaf*. Retrieved May 5, 2006, from http://www.signingbooks.org/

Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *Journal of Technology, Learning, and Assessment*, 2(6). Available from http://www.jtla.org

Pragma, NL. (2002). *Sign Language Animations*. Available at http://www.signingbooks.org/animations/sign_language_animations.htm

Richards, J.C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly*, 17 , 219-240.

Rixon, S. (1981). *The design of materials to foster particular linguistic skills. The teaching of listening comprehension*. (ERIC Document Reproduction Service No. ED 258 465).

Rose, D. & Meyer, A. (2002). *Teaching Every Student in the Digital Age: Universal Design for Learning*. Alexandria, VA.: Association for Supervision and Curriculum Development.

Rose, D. & Dalton, B. (2002) Using Technology to Individualize Reading Instruction. In Block, C. Gambrell, L. & Pressley, M. *Improving Comprehension Instruction*, Jossey Bass, N.Y., N.Y>

Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 15-30). Englewood Cliffs, NJ: Prentice-Hall.

Rubin, J. (1995). The contribution of video to the development of competence in listening. In D.J. Mendelsohn & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 151-165). San Diego, CA: Dominie Press.

Safran, J. S. (2002). Supporting students with Asperger's Syndrome in general education. *Teaching Exceptional Children*, 34(5), 60-66.

Schunk, D.H. & Zimmerman, B. J. (1997). Developing self-efficacious readers and writers: The role of social and self-regulatory processes. In J.T. Guthrie & A. Wigfield (Eds.) *Reading engagement: Motivating readers through integrated instruction* (pp.34-50). Newark, DE: International Reading Association.

Science Accessibility Project (2003). *Audio Accessible Graphing Calculator*. Corvallis, OR: Oregon State University. Demo available at http://dots.physics.orst.edu/calculator.html

Smisko, A., Twing, J., & Denny, P. (2000). The Texas model for content and curricular validity. *Applied Measurement in Education, 12*(4), 333-342.

Southwest Educational Development Laboratory. (2004). *Accessibility Design Considerations and Examples*. Retrieved May 5, 2006, from http://www.ncddr.org/du/researchexchange/v02n01/design.html

Stahl, S. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension: The retiring of number 41. *Journal of Reading Behavior* 23(4), 487-508. Chicago, ILl: National Reading Conference.

Strangman, N., & Dalton, B. (2005). Technology for struggling readers: A review of the research. In D. Edyburn et al (Eds). *The Handbook of Special Education Technology Research and Practice* (pp.545-569). Whitefish Bay, WI: Knowledge by Design.

Stone, C.A. & May, A.L. (2002). The accuracy of academic self-evaluations in adolescents with learning disabilities. *Journal of Learning Disabilities*, 35 (4) 370-83.

Thompson, A. D.; Sproule, S.L. (2000). Deciding when to use calculators. *Mathematics Teaching in the Middle School 6*(2), 126-129.

Trace Research and Development Center. (1994). *Application Software Design Guidelines: Increasing the Accessibility of Application Software to People with Disabilities and Older Users*. Madison, WI: University of Wisconsin-Madison. Available at: http://trace.wisc.edu/docs/software_guidelines/toc.htm

Underwood, M. (1989). *Teaching listening*. London: Longman.

Unok-Marks, S., Shaw-Hegwer, J., Schrader, C., Longaker, T., Peters, I., Powers, F., et al. (2003). Instructional management tips for teachers of students with autism spectrum disorder (ASD). *Teaching Exceptional Children*, 35(4), 50-55.

Vcom3D Inc. (2006). *Vcom3D*. Retrieved May 5, 2006, from http://www.vcom3d.com

View Plus Technology (2006). *Tiger Software Suite*. Information available at http://www.viewplustech.com

Vischeck.com. (n.d.). *Color Blindness Accessibility* tool. Available at http://www.vischeck.com/vischeck/vischeckURL.php.

W3C-WAI. (1999). *Web Content Accessibility Guidelines 1.0*. Retrieved May 5, 2006, from http://www.w3.org/TR/WCAG10/

Wang, J.H.Y. & Guthrie, J.T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly*, 39, 162-186.

Wood, K.D. (2002) Differentiating reading and writing lessons to promote content learning. In C. Collins Block, L. B. Gambrell, M. Pressley (Eds.) *Improving comprehension instruction: Rethinking research, theory, and classroom practice*. pp. 155-180. San Francisoc: Jossey-Bass.

World Wide Web Consortium. (2). *User Agent Accessibility Guidelines 1.0*. Available at: http://www.w3.org/TR/UAAG10/

World Wide Web Consortium. (3). *Scalable Vector Graphics Specifications*. Available at: http://www.w3.org/TR/SVG/

Ziomek, R. & Andrews, K. (1996). *Predicting the College Grade Point Averages of Special-tested Students from their ACT Assessment Scores and High School Grades*. Iowa City, IA: ACT.

Appendix: Checklists

Text

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Refreshable Braille, Screen Reader, TTS
	Visual Acuity	Flexible size text
	Visual Discrimination	Flexible fonts, flexible contrast
Linguistic	English Language Proficiency	Alternate languages (natural, ASL, non-English)
	Vocabulary	Vocabulary links (dictionary & thesaurus, talking, multiple languages)
	Syntactic Skills	Grammar aids, simplified syntax
	Word Decoding Skills	TTS for individual words (talking dictionary)
	Reading Fluency	TTS with synchronous highlighting
	Knowledge of Text Structure	Graphic organizers, explicit indicators of text structure
Cognitive	Background Knowledge	Links to background knowledge
	Comprehension Strategies	Prompts and supports for strategies
	Categorical and Conceptual Skills	Advance organizers, concept maps
	Attention and Concentration skills	Prompting, breaking text into smaller sections, locate items near relevant text
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-Specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Images

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Tactile display, 3-d manipulatives, text equivalents, longdesc
	Visual Acuity	Flexible image size, zoom
	Visual Discrimination	Flexible contrast
	Color Perception	User specified color options, avoid common color-blindness combinations, redundant presentation of information conveyed in color
	Shape Recognition	Alternative visual options, description, tactile option
Cognitive	Visual Processing Skills	Highlight critical features
	Knowledge of Graphic Conventions	Alternative descriptions or depictions
	Knowledge of Iconic Conventions	Alternatives for icons (rollover descriptions, legend, customizable icons for tools or commands)
	Visual Syntax Fluency	Highlight critical relationships
	Background Knowledge	Links to background knowledge
	Cognitive Strategies	Prompts and supports for viewing and interpretation strategies
	Planning and Organizing Skills	Graphic organizer, planning templates
	Attention and Concentration	Simplified images, prompts
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Audio

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Hearing Ability	Visual alerts, captions (SMIL, etc.)
	Auditory Threshold	Adjustable volume
	Auditory Processing Speed	Adjustable rate
	Auditory Discrimination	Highlight critical features, emphasize discriminants
Linguistic	English Language Proficiency	Alternate languages (natural, ASL, non-English)
	Receptive Vocabulary	Vocabulary links to predetermined lists or dictionary with word prediction for spelling
	Syntactic skills	Alternate syntactic levels (simplified text)
	Prosody Recognition	Alternative prosodic emphasis
	Idiomatic Expression Familiarity	Alternatives for idiomatic expressions
Cognitive	Background Knowledge	Links to background knowledge
	Cognitive Listening Skills	Prompts and supports for listening and interpretation strategies
	Planning and Organizing Skills	Graphic organizer, planning templates
	Attention and Concentration	Increased segmentation, navigation control (pause, forward, reverse, replay, and search features
Motoric	Navigation of Audio File	Keyboard alternatives for all on screen navigation commands, assistive device compatibility, do not disable OS functions
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions; review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Tables and Graphs

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Identify row and column headers using appropriate mark up language; identify subheads and data cells using appropriate mark up language; provide a linearized version for Screen Readers that cannot read tables or side by side text; do not use structural markup for visual formatting, it interferes with screen readers; provide summaries of tables for VI; static tables can be rendered with pre-produced audio (MathSpeak, National Braille Association Tape Recording Manual)
	Visual Acuity	Flexible size, zoom, SVG or similar for static tables/graphs
	Visual Discrimination	Flexible Contrast
	Color Perception	User specified color options, avoid common color-blindness combinations, redundant presentation of information conveyed in color
	Display Complexity	Present only necessary information; explicit labeling and formatting
Cognitive	Knowledge of Conventions	Highlight critical features, Alternative descriptions or depictions
	Visual syntax Fluency	Highlight critical relationships
	Background Knowledge	Links to background knowledge
	Cognitive Strategies	Prompts and supports for viewing and interpretation strategies
	Planning and Organizing Skills	Graphic organizer, planning templates
	Attention and Concentration	Tools for orientation within table/graph
Motoric	Navigating Tables and Graphs (Static and Dynamic	Keyboard alternatives for all on screen navigation commands, long desc of commands for active tables or graphs including result of action; assistive device compatibility, do not disable OS functions
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant screen elements
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Mathematical and Scientific Notation

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Nemeth Code, MathML (www.w3c.org/Math) LaTeX, ChemML, CML (http://www.xml-cml.org/information/position.html), AsTeR (http://www.cs.cornell.edu/Info/People/raman/aster/aster-toplevel.html); flexible size, pictorial representations
	Visual Acuity	Flexible Fonts, Zoom, SVG or similar technology
	Visual Discrimination	Flexible Contrast
Linguistic	Mathematical Syntax	Highlight order of operations
	Mathematical Fluency	Simplified numbers, retain concept
Cognitive	Background Knowledge	Links to background knowledge
	Calculations Complex	Calculator, scrap paper, simplified numbers
	Expressions	Make expressions accessible to screen readers with text descriptions using MathML or LaTeX, following guides for spoken mathematics (NCAM). Embed text and audio files, gets tricky with complex expressions because each part needs to be accessible separately as well wholly. Concatenated speech is awkward for complex expressions. AsTeR reads LaTeX, creates audio, and allows navigation (http://www.cs.cornell.edu/Info/People/raman/aster/aster-toplevel.html)
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Video/Animation

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Rich description
	Visual Acuity	Flexible image size, zoom
	Visual Discrimination	Black and White/Greyscale options, flexible contrast
Cognitive	Visual Processing Skills	Highlight critical features
	Knowledge of Graphic Conventions	Alternative descriptions or depictions
	Knowledge of Iconic Conventions	Alternatives for icons
	Visual Syntax Fluency	Highlight critical relationships
	Background Knowledge	Links to background knowledge
	Cognitive Strategies	Prompts and supports for viewing and interpretation strategies
	Planning and Organizing Skills	Graphic organizer, planning templates
	Attention and Concentration	Summary of action, prompts, navigation control (pause, forward, reverse, replay and search features), highlighting in graphic organizer synchronized to stages of action
Motoric	Navigation of Animation or Video	Keyboard alternatives for all on screen navigation commands; assistive device compatibility, do not disable OS functions
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Response Options

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Ability to Distinguish Stimulus and Response Components	Explicit labeling; physical and functional separation of stimulus and response areas, with supports for navigation between them by screen readers and single switch devices
	Ability to Ascertain Actions Required for Response	Simple, clear instructions; highlight all enabled elements; consistent methodology for highlighting enabled elements; simultaneous highlighting of functionally related enabled elements; clear relationship between mouse active behaviors and mouse keys, tab navigation, single switch navigation; animations to model required actions
Linguistic	English Language Proficiency	Alternate languages (natural, ASL, non-English),
	Vocabulary	Vocabulary links (dictionary & thesaurus, talking, multiple languages)
	Syntactic Skills	Grammar aids, simplified syntax
	Word Decoding Skills	TTS for individual words (talking dictionary)
	Reading Fluency	TTS with synchronous highlighting
	Knowledge of Text Structure	Graphic organizers, explicit indicators of text structure
Cognitive	Understanding Response Requirements	Animation of tools required for response, practice test training, consistent tool use across items
	Planning and Organizing Skills	Graphic organizer, planning templates
	Attention and Concentration	Divide task into discrete steps
Motoric	Navigation Abilities	Do not disable OS functions to ensure assistive device and software compatibility; keyboard alternatives for all on screen and mouse active commands, tab navigation, Voice activation
	Selection	Dwell time selection Assigned key (tab, space)
	Keyboarding	Alternate Keyboard, Screen Keyboard (fatiguing), Dictation (Scribe or Voice Recognition)
	Drag and Drop	Assigned keys for select, hold, drop, Keyboard equivalents, Structured navigation with tabs, , Snap to constraints
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Active Objects and Links

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Text equivalents for all non-text elements
	Visual Acuity	Flexible image size, zoom
	Visual Discrimination	User specified color options, avoid common color-blindness combinations, redundant presentation of information conveyed in color, flexible contrast
	Shape Recognition	Alternative visual options, description
Cognitive	Visual Processing Skills	Redundant text links, client side image maps, hot spots should be rendered as list of hypertext links
	Knowledge of Active Object Conventions/ Hypertext Syntax Fluency	Explicit instructions, semantic info of objects conveyed through text, objects that represent controls/tools or other programmatic elements must be used consistently throughout, clearly indicate the target of all links, highlight critical relationships (color or highlighting to indicate related enabled elements, redundant text or auditory indicators for VI)
	Attention and Concentration	Prompts, explicit descriptions, focus indicator
Motoric	Navigation Abilities	Keyboard alternatives for all on screen navigation or action; assistive device compatibility, do not disable OS functions
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Constructed Response: Text Composition

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Alternate Input Devices, Braille displays, TTS or self-voicing read-back of student composition
	Visual Acuity	Flexible size text
	Visual Discrimination	Flexible fonts, flexible contrast
Linguistic	English Language Proficiency	Alternate languages for composition (natural, ASL, non-English), TTS or self-voicing read-back
	Vocabulary	Vocabulary links (dictionary & thesaurus, talking, multiple languages, ASL translator)
	Syntactic skills	Grammar check, TTS or self-voicing read-back
Cognitive	Medium Familiarity/Dexterity	Alternate response options
	Planning and Organizing Writing	Graphic organizers, access to rubric
	Writing Fluency	Models, virtual mentors
Motoric	Production Dexterity	Alternate input devices, dictation (voice recognition, scribe), do not override OS functions
	Navigation Abilities	Keyboard alternatives, assistive device compatibility (do not override OS functions)
	Strength and Mobility	Assistive Device Compatibility (do not override OS functions), physical setting flexibility
	Automaticity	Variable or no time constraints
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions; review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Constructed Response: Math

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Visual Ability	Alternate Input Devices, Braille displays, particularly complex for math to be read-back, complex for students entering Nemeth Code
	Visual Acuity	Flexible size font
	Visual Discrimination	Flexible fonts, flexible contrast
Linguistic	Creating Graphs or Tables	Templates, edit or reorganize vs. create, spreadsheets, graph + scrap paper
	Creating Diagrams or Drawings	Drawing palettes
	Creating Equations	Equation palettes, on-screen calculators
Cognitive	Calculations	On screen calculators, simplified calculations
	Problem Solving	Models, virtual mentors, calculation focus
Motoric	Production Dexterity	Alternate input devices, dictation (voice recognition, scribe), do not override OS functions
	Navigation Abilities	Keyboard alternatives, assistive device compatibility (do not override OS functions)
	Strength and Mobility	Assistive Device Compatibility (do not override OS functions), physical setting flexibility
	Automaticity	Variable or no time constraints
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content

Multi-stage/Multi-part Items

Category of Processing	If this item *does not* intend to measure:	Then consider the following design options to minimize measurement of unintended constructs (construct irrelevant variance):
Perceptual	Distinguishing Intra-item Navigation Actions	Consistent highlighting of all enabled intra-element navigation indication and control elements; clear relationship between mouse active behaviors and mouse keys, tab navigation, single switch navigation
	Identifying Stimulus and Response Components	Consistent, distinguishable stimulus and response designs; explicit labeling; physical and functional separation of stimulus and response areas, with supports for navigation between them by screen readers and single switch devices
Linguistic	Constructing Meaning From Text	Provide clear concise description of stages and parts; provide alternate representations of the stages and parts
Cognitive	Understanding Response Requirements	Provide explicit instruction to indicate required steps; use elements consistently; provide practice; be consistent with instructional practices
	Planning and Organizing Skills	Allow actions to be reversed; allow reverse navigation; allow modification of previous stages and parts; allow selective clearing of work
	Hypertext Syntax Fluency	Use alt-text; indicate link targets
Motoric	Navigation Abilities	Provide keyboard alternatives for intra-item navigation and mouse actions; ensure sensible tabbed navigation; do not disable OS functions to ensure assistive device and software compatibility
Executive	Goal Setting Ability	Explicit Instructions, goal-setting supports
	Goal Maintenance and Adjustment	Reminders, prompts
	Monitoring Progress	Extrinsic Scaffolds for monitoring
	Working Memory	Note-taking, mnemonic aids, text complement, locate items near relevant text
Affective	Self-regulation	Scaffolds for self-regulation
	Intrinsic Task-Specific Motivation	Alternative content for interest
	Extrinsic Incentives	Individualized rewards, repercussions
	Test Conditions	Alternative settings and conditions (time, sessions, location); review for racial, cultural, ethnic, & gender bias; differential item functioning; age appropriate content