Defining Reading Proficiency for Accessible Large-Scale Assessments: Some Guiding Principles and Issues

NARAP

National Accessible Reading Assessment Projects

National Accessible Reading Assessment Projects
http://www.narap.info

Partnership for Accessible Reading Assessment
University of Minnesota (NCEO and Dept. of C & I) / CRESST (UCLA & UCD) / Westat
350 Elliott Hall
75 E. River Road
Minneapolis, MN 55455
(612) 625-7241
Fax: (612) 624-0879
http://www.readingassessment.info
readingassess@umn.edu

Designing Accessible Reading Assessments
Educational Testing Service
Rosedale Road, MS 09R
Princeton, NJ 08541
(609) 734-1347
Fax: (609) 734-1755
http://www.ets.org/dara
readingassessment@ets.org

This work is supported, in part, by the U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research—Grant Numbers H324F040001 and H324F040002. Opinions expressed do not necessarily reflect those of the U.S. Department of Education or offices within it. Endorsement by the Federal Government should not be assumed.

Defining Reading Proficiency for Accessible Large-Scale Assessments: Some Guiding Principles and Issues

Prepared by:

National Accessible Reading Assessment Projects

Designing Accessible Reading Assessments (DARA)
Educational Testing Service (ETS)

and

Partnership for Accessible Reading Assessment (PARA)
University of Minnesota—NCEO (National Center on Educational Outcomes) and Department of Curriculum and Instruction

CRESST (Center for Research on Evaluation, Standards, and Student Testing—University of California, Los Angeles; and University of California, Davis)

Westat

February 17, 2006

Please use the following citation for this paper:

National Accessible Reading Assessment Projects. (2006). *Defining reading proficiency for accessible large-scale assessments: Some guiding principles and issues.* Minneapolis, MN: Author.

Available on the World Wide Web at www.narap.info.

Introduction

The National Accessible Reading Assessment Projects (NARAP) is a collaboration of two projects funded by the U.S. Department of Education to conduct research and development on accessible reading assessments for students with disabilities that affect reading. The goal of these projects is to produce research findings and assessment techniques that demonstrate how large-scale assessments of reading proficiency can become more accessible and valid for all students, while also meeting the assessment requirements of the No Child Left Behind Act of 2001[1] (NCLB).

Understanding the need for accessible assessments is key to determining the steps that must be taken to develop such assessments. NCLB's critical provision that all students be included in state accountability systems, including the full range of students with disabilities, brings to the forefront assessment challenges that have not yet been adequately addressed. To be fair to schools and to students, assessments need to allow all students the best opportunity to show what they know and what they can do. All students have skills and knowledge in reading that reflect varying levels of proficiency by reading component. That is, each student may have some reading skills and knowledge that reflect what is considered to be the proficient level, other skills and knowledge that are emerging, and other reading skills and knowledge that have not yet emerged. Some students with disabilities may have skills and knowledge that are expressed in ways that are not measured in current assessments; that is, they may have skills and knowledge that in many assessments are integrated with skills and knowledge they do not have (e.g., when decoding of passages is required in order to demonstrate comprehension). Thus, the definition of reading proficiency must be stated in such a way that allows assessments to reflect what students are able to do, not just what students are not able to do.

The desire for all students with disabilities to be able to show their proficiency as readers does not indicate a desire to lower expectations for the proficiency levels of these students. Rather, it is important for students to be held to the same standards of reading proficiency whenever possible. Still, it will likely be necessary to provide flexible expressions of reading. Flexible expressions of reading may include accommodations such as letting students give oral responses to questions, but may also require an approach that is built into the assessment itself – universal design. By universal design we mean assessments that are designed from the beginning, and continually refined, to allow participation of the widest possible range of students, resulting in more valid inferences about performance (Thompson, Johnstone, & Thurlow, 2002). Technology may be an element of providing universal design and flexible access to all students, including students with disabilities.

Seeking flexibility also may suggest the need for large-scale assessments that do not rely solely on one way for students to demonstrate proficiency. Students who are able to demonstrate skills on a range of important reading standards may be considered readers even though they may not be proficient on all components. Alternative approaches

1 This includes using the definition of reading that is included in the Reading First program of NCLB.

to scoring assessments – such as the use of compensatory approaches rather than conjunctive approaches – may be another avenue to realizing a definition of reading proficiency. By this we mean, for example, being able examine comprehension skills separately from decoding skills. These alternative approaches recognize that students with disabilities deserve the opportunity to show the full profile of their developing proficiency as readers.

Consequently, one of the first steps in designing accessible assessments is the precise definition of constructs being measured. Such definitions permit construct-relevant factors to be distinguished from construct-irrelevant factors, thus facilitating the removal of barriers to accessibility while maintaining valid measurement of the constructs of interest. Similarly, distinguishing various components of the construct may facilitate development of assessments that can reveal strengths in some components that would otherwise be masked by weaknesses in other components. Thus, one of NARAP's first tasks was to develop a definition of "reading proficiency." In order to accomplish this we collected information in a variety of ways, including: (a) reviewing existing definitions of reading proficiency (e.g., various reports about reading including the NRP, RAND, PISA, PIRLS), (b) convening a panel of experts to provide input, and (c) conducting focus groups. (The information and materials used during this process as well as the focus group findings are available at www.narap.info .) This paper is a synthesis of the information collected and describes a set of principles and key unresolved issues described under each principle. These principles and issues will guide the next phases of our project (research and development). This paper does not focus on the assessment itself; rather, the assessment developed will be informed by our research and then addressed in a separate paper focusing on the principles and guidelines of accessible assessments.

Principles for Defining Reading Proficiency

NARAP has identified three principles for defining reading proficiency. Each of these principles is discussed here, with issues identified as appropriate.

Principle 1: Definitions of reading proficiency must be consistent with core NCLB provisions.

The federal priority under which NARAP was funded cited two required purposes of NCLB assessments that must be met by accessible large-scale tests of reading proficiency. These purposes are to "provide (a) a valid measure of proficiency against academic standards, and (b) individual interpretive, descriptive, and diagnostic reports for the full range of students with disabilities that affect reading." With regard to the first of these two purposes, NCLB requires each state to establish grade-level academic standards. Because each state sets its own standards, the definition of reading proficiency used by NARAP must not impose a particular standard but rather allow for variability among states. To be consistent with NCLB and its requirement for grade-level standards,

NARAP must base its research and development on grade-level definitions of reading proficiency. Access, participation in, and progress in the general curriculum is a foundational requirement of the Individual with Disabilities Education Act (IDEA) since 1997, reinforced as access, participation, and progress in grade-level curriculum and standards by NCLB law and regulation and IDEA 2004 law and draft regulations.

States address areas of reading proficiency and component skills through grade-level definitions of content standards (Thompson, Johnstone, Thurlow, & Clapper, 2004). For example, phonemic knowledge is on the standards lists of 47 states at grades 1 and 2, and 27 states at grade 8. Fluency standards are listed for 39 states in grades 1 and 2 and 30 states at grade 8. Skill standards for expository elements of text are listed by 20 states at grades 1 and 2, 33 states at grade 4, and 30 states at grade 8. The number of states listing the same standards for all grades (1-8) on other aspects of reading skills is fairly consistent (e.g., inferential comprehension: grades 1 and 2—on the lists of 31 states, grade 4—on the lists of 37 states, grade 8—on the lists of 40 states and word recognition is on the lists for 45 states for grades 1 and 2 and remains on the lists of 40 states for grade 8). Such general statements, however, should be read with caution, since the method of simply counting references to terms in standards may not capture the relative emphasis of standards in particular grades. There are additional caveats to this analysis. First, an analysis of state standards does not necessarily indicate the competencies states actually include on their assessments. Second, current practice in defining state reading standards may not be best practice in some cases. In fact, some states do not measure appropriate skills and are conservative about what skills are listed or measured.

Grade-level skills are important and necessary foundations for any accessible assessment. This is evident in NAEP, which all states now are required to administer. The NAEP 2009 framework provides a model for defining reading proficiency in grades 4, 8, and 12 by requiring students to read both literary (fiction, nonfiction, and poetry) and information texts (exposition, argumentation, persuasive, and procedural texts or documents). Readers must access words in texts, use the structure of texts, make sense of vocabulary as it is embedded in a text, understand sentences and paragraphs, and comprehend what they read. NAEP assessments expect that students' text comprehension will be influenced by their ability to apply the foundational components of reading: phonemic awareness, phonics knowledge, fluency, and vocabulary. In addition, NAEP 2009 focuses on grade-level "cognitive targets" and defines these as "the mental processes or kinds of thinking that underlie reading comprehension; the cognitive targets serve to guide the test development process in that item writers 'target' these processes or kinds of thinking as they write items" (2009 NAEP Reading Framework, p. 39).

The second aspect of NCLB cited above, that assessments should provide individual interpretive, descriptive, and diagnostic reports for the full range of students with disabilities that affect reading, emphasizes that assessments should be as worthwhile as possible for the full range of students. To the extent that this is possible, assessments used primarily for accountability purposes (e.g., NCLB), should also provide useful information to educators as they plan instructional improvements. This information

would be coupled with other reading assessments used at a local level that are individually administered and diagnostic in nature (e.g., informal reading inventories; running records). Literature on developing and using tests provides cautions about the extent to which a single test can be used for multiple purposes (AERA, APA, NCME, 1999).

Issues related to Principle 1:

A. How do the important reading skills vary as a function of grade level?

B. How do we determine which measure may be appropriate for use at a specific grade level?

C. How much can achievement levels vary and still meet the requirements of grade-level content?

D. How are differences in reading achievement standards (e.g., modified or alternate achievement standards) developed and defined? How are these varying achievement standards reflected in definitions of reading proficiency?

Principle 2: Reading proficiency must be defined in such a way that flexible expressions of reading are allowed while preserving the essential nature of reading. This is crucial as we seek to make assessments accessible to students with a variety of disabilities.

Flexibility in how reading proficiency is defined and measured is demanded by the practice of states developing their own standards and the reality that students differ in their reading strengths and weaknesses in component processes. Proficient readers draw upon a range of processing abilities ranging from lower-level processes to higher-order processes. At the lower level they recognize individual words, and at higher levels they assemble information from multiple sources into meaningful representations of text and relate it to background knowledge. No one component process can account for overall proficiency. Rather, both lower-level and higher-level processes contribute to individual differences in reading proficiency (Daneman,1996). The reliance on any set of component proficiencies and the use of compensations are aimed at allowing readers to achieve overall proficiency in understanding a given text. Typical proficient readers rapidly recognize words thus freeing up processing capacity for comprehension. On the other hand, readers who have visual or auditory disabilities may compensate for a lack of rapid recognition of orthographic patterns through more efficient memory of written discourse structure for expository texts or stories than other readers. A reader using braille can demonstrate overall reading comprehension proficiency, albeit by means of an alternate format. Students with congenital deafness may achieve overall reading comprehension proficiency, but may need to deploy alternative strategies to understanding sound-symbol relationships as a basis for reading.

Public expectations and perceptions about reading often are reflected in practice and public policy. We live in a world of printed text, but also a world with growing availability of text in other formats, such as books on tape or CD, printed text translated into speech by reading machines, text transmitted electronically via the internet and readable by screen reading software, and materials in "accessible" electronic formats (e.g. digital talking book and the National Instructional Materials Accessibility Standard) that can be presented to the reader in various visual or auditory modalities. Individuals who are blind or partially sighted often refer to their activity as "reading" when they are in fact listening to text, and people without disabilities also take advantage of these new formats, for example by listening to books while engaging in other activities.

Yet, the public is more likely to view the individual who is "reading" braille as a "reader" than they are to refer to the individual who is "reading" American Sign Language (ASL) as a "reader." Similarly, the reader of ASL is more likely to be considered a reader than the person who is "reading" by having a screen reader read a page of text to him or her. These public views are reflected in the number of states that allow these types of approaches on their state assessments – braille is allowed without restriction in 39 states; sign language interpretation of the reading test is allowed without restriction in 13 states; and reading the questions aloud to the test taker is allowed without restriction in 3 states (Clapper, Morse, Lazarus, Thompson, & Thurlow, 2005).

Federal statutes such as NCLB and IDEA allow a range of options in the types of assessments used and the achievement standards applied to students with disabilities. In addition to taking general assessments based on grade-level achievement standards with or without accommodations, some students with disabilities can take assessments based on modified achievement standards, or alternate assessments based on grade level standards or (for students with the most significant cognitive disabilities) on alternate achievement standards. Modified or alternate achievement standards must be aligned to grade-level content standards, but may differ from the grade-level achievement standards in breadth, depth, or complexity. All of these are possible means of flexibility for students with disabilities.

Issues related to Principle 2:

A. How broadly can we define what constitutes "reading" and still have the definition based on grade-level achievement standards?

B. Can what constitutes "reading" for standards-based assessments differ by disability category or by needed accommodation?

C. How do the concepts of modified and alternate achievement standards apply to grade level reading standards?

D. How do students with disabilities compensate for weaknesses in specific reading proficiency components due to their disability or multiple disabilities?

Principle 3: Definitions of reading proficiency must reflect both comprehension and foundational skills.

The information we collected by reviewing existing definitions of reading proficiency, obtaining input from a panel of experts and conducting focus groups helped develop a firm base upon which to anchor the flexibility described in Principle 2. Although the emphasis varied, a consistent message was that the core construct of a definition of reading references both foundational reading skills and comprehension. The following definition of reading, included in the *Reading First* program of *No Child Left Behind* (2001), places a clear emphasis on foundational reading skills.

The term reading means a complex system of deriving meaning from print that requires all of the following:

(A) The skills and knowledge to understand how phonemes, or speech sounds, are connected to print.

(B) The ability to decode unfamiliar words.

(C) The ability to read fluently.

(D) Sufficient background information and vocabulary to foster reading comprehension.

(E) The development of appropriate active strategies to construct meaning from print.

(F) The development and maintenance of a motivation to read.

Other definitions such as the 2009 NAEP framework stress that "reading is an active and complex process that involves understanding written text; developing and interpreting meaning; and using meaning as appropriate to type of text, purpose, and situation." In this definition it is assumed that students must apply foundational skills to comprehend a variety of texts. Thus, an assessment based on the NAEP framework is a meaningful indicator of reading comprehension that indirectly measures foundational skills. It could be argued that if a child can demonstrate proficiency on such an assessment, no additional assessment of foundational skills is required. On the other hand, defining the foundational skills involved in the reading processes is particularly important for the population under study by NARAP (students with disabilities that affect reading) because we cannot assume that these foundational skills have been or are in the process of being acquired. Defining foundational skills will give NARAP the flexibility to develop component level measures that provide information on what students can do, rather than a single score of non-proficient. This principle suggests the need for large-scale assessments that are flexible (perhaps technology-based) and able to assess reading comprehension and/or foundational skills based on student performance.

Issues related to Principle 3:

A. To what degree can component skills be measured independently?

B. Comprehension is the primary goal for readers. If students are proficient in this area with accommodations, do we need to measure the foundational skills?

C. If comprehension is our primary goal, should the comprehension score be weighted more heavily when foundational skills are also assessed?

D. What feasible techniques are available for measuring foundational skills such as fluency or phonics knowledge in the context of large-scale assessment?

E. For students who do not achieve grade-level proficiency, what processes can be developed or applied to aggregate their performance on component skills into an overall measure of reading proficiency?

F. Some components appear to be problematic for certain disabilities (e.g. phonemes and deaf students). Do we develop alternate definitions of proficient reading for these populations based on a better understanding of reading processes and performances?

G. Are some skills less critical to measure than others?

H. If foundational skills were only assessed after a student had performed below proficient on comprehension, what proportion of ALL students (with and without disabilities) would be assessed on each of the foundational skills? Is this proportion small enough to assess students in small groups, individually, or via computer?

I. Can some foundational skills be assessed together (e.g., decoding and phonemes)?

J. If foundational skills are going to be measured only for students who are not proficient on an assessment of comprehension, can accommodations be allowed that invalidate the foundational skills (e.g., read aloud for decoding or extra time for fluency)?

K. If a student is not proficient on a measure of reading comprehension, should listening comprehension be assessed prior to measuring foundational skills?

Conclusion

This paper focuses on several of the complex issues in creating accessible and valid measures of reading proficiency for students with disabilities that affect reading. The principles included in this paper will guide NARAP in formulating the definitions of reading proficiency we will use within our studies. A definition of reading proficiency, compatible with all three principles outlined in this paper, would be an organizing structure or framework that would support each state as educators indicate how much emphasis various components of reading have in their grade-level reading standards. During the next phase of our project we will focus our efforts on conducting research that will be consistent with Principle 1 and address some of the issues listed under principles 2 and 3. Several of these issues cannot be resolved through empirical research. Instead, they require that we focus our efforts on providing information that can be used to set policies on assessing students with disabilities that impact reading.

References

American Educational Research Association, American Psychological Association, and National Council of Measurement in Education (1999). *Standards for educational and psychological testing.*

Clapper, A., Morse, A., Lazarus, S., Thompson, S., & Thurlow, M. (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.

Daneman, M. (1996). Individual differences in reading skills. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (vol. 2) (pp. 512–538). Mahwah, NJ: Erlbaum.

Individual with Disabilities Education ACT of 1997, 20 U.S.C. 1412 (a) (17) (A) (1997).

National Assessment Governing Board. (Spring, 2005). *Specifications for the 2009 NAEP Reading Assessment.* Washington, DC: Author. Retrieved February 18, 2006, from http://www.nagb.org/pubs/reading_fw_08_05_prepub_edition.doc
[Also, see Salinger, T., Kamil, M. L., Kapinus, B., & Afflerbach, P. (2005). Development of a new framework for the NAEP reading assessment. *54rd Yearbook of the national reading conference (pp. 334–348)*. Oak Creek, WI: National Reading Conference.]

No Child Left Behind Act of 2001, 20 U.S.C. 6301 e seq (2001) PL 107–110).

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thompson, S. J., Johnstone, C. J., Thurlow, M. L., & Clapper, A. T. (2004). *State literacy standards, practice, and testing: Exploring accessibility* (Technical Report 38). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.