

White Paper

Building Reading Benchmark Assessments that Align with Common Core State Standards

by
Lucas J. Schippers, Ph.D.



**Assessment
Technology
Incorporated**

Assessment Technology, Incorporated

6700 E. Speedway Boulevard
Tucson, Arizona 85710

Phone: 520.323.9033 • Fax: 520.323.9139

Copyright © Assessment Technology, Incorporated 2012. All rights reserved.

Copyright © Assessment Technology, Incorporated 2012. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the publisher.

Assessment Technology, Incorporated, Publishers

Tucson, Arizona, U.S.A.

Printed in the United States of America.

V2-071112

Building Reading Benchmarks Assessments that Align with Common Core State Standards

*By Lucas J. Schippers, Ph.D.
Assessment Technology, Incorporated*

Table of Contents

| | |
|---|----------|
| Table of Contents..... | i |
| I. Introduction | 1 |
| II. New Standards, New Assessments | 1 |
| III. Reading in Common Core | 2 |
| IV. Building Valid Common Core State Standard Reading Assessments..... | 2 |
| V. Discussion | 3 |
| VI. References..... | 4 |

This page intentionally left blank.

I. Introduction

With the implementation of Common Core State Standards (CCSS) currently underway, educators are under pressure to respond to a new set of educational expectations. This concern is particularly acute vis-à-vis assessment, given the accountability systems in use today.

This paper offers guidance to educators who want to ensure that their benchmark assessments of reading effectively measure CCSS. The following questions will be addressed:

1. Why do new standards demand changes to assessment?
2. What expectations of student readers are present in CCSS?
3. How should reading benchmark assessments be constructed to reflect these expectations?

II. New Standards, New Assessments

Why should new standards engender changes to an assessment program? Simply put, assessments that do not evolve along with the standards they purport to measure are of questionable validity. According to the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, validity is “the degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses” (1999). An assessment program can be thought of as an argument in support of the claim that the assessments are valid for their intended purpose, and for which test scores are collected as evidence; if the argument is logical and reasonable and is supported by evidence, the argument can be said to be valid (Haladyna, 2004).

A set of standards typically establishes a body of content knowledge that students are expected to learn. When standards change, so too does the prescribed content knowledge. Updating assessments to ensure coverage of new content helps guard against loss of content validity (i.e., the comprehensiveness of a test’s measurement of content). However, standards are more than the sum of the content knowledge they prescribe; undergirding a set of standards is a theory of the discipline (e.g., what it means and entails to be a reader). When this theory is operationalized in curriculum and pedagogy, it becomes a construct, and assessments that accurately reflect the construct possess construct validity. The format of a test, including the types of items used, influences construct validity (Messick, 1989). Moreover, construct validity is enhanced when the assessment represents disciplinary processes in a way that reflects the theory of the content (e.g., students read and analyze texts included on a test in accordance with the processes sanctioned by the standards) (Dunbar, Koretz, & Hoover, 1991).

In sum, care must be taken in the construction of assessments to ensure that they measure what they are said to measure. The adoption of a new set of standards brings into sharp relief concerns over an assessment’s content and construct validity. Updating the content that is tested as well as the way it is tested helps preserve the validity and usefulness of assessment data.

III. Reading in Common Core

Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects (2010) develop a construct of reading characterized by the following assumptions:

- **Students should read more challenging texts.** The authors of the standards argue that texts read in schools are not complex and demanding enough to prepare students for the kinds of materials usually encountered in college courses and the workplace. Thus, the standards call for students to engage with increasingly complex texts as they proceed through the grades.
- **Students should read more attentively and respond more deeply to texts.** The complexity and depth of analysis mandated by the standards require students to spend significant time and energy reading texts. Partnership for Assessment of Readiness for College and Careers (PARCC), one of two consortia of states working to develop CCSS assessments, recommends that language arts curricula developed for CCSS emphasize “close, sustained analysis of complex text” so that students can achieve “understanding of the text as a whole” (2011, p. 6). Thus, the quality of students’ reading and analysis is prioritized over shallow reading of a greater number of texts. PARCC recommends teaching one long text (e.g., a novel) per quarter along with an assortment of shorter, thematically complementary texts (e.g., essays).
- **Students should read more informational text.** As per the ratios established by the National Assessment of Educational Progress, the standards give significant weight to reading informational text. Students are expected to read informational text in equal measure to literary text in the elementary grades, while the ratio shifts to 70/30 in favor of informational text for high school students. Moreover, high school students are expected to read far more literary nonfiction (e.g., Patrick Henry’s “Speech to the Second Virginia Convention”), a kind of informational text, than has traditionally been assigned.
- **Students should write about what they read.** Even though reading and writing are addressed in separate sections in the document, the standards assume that the processes of literacy are integrated in practice. For example, the writing standards directly call upon students to apply in writing what they have learned from their analysis of a text.

IV. Building Valid Common Core State Standard Reading Assessments

The above analysis suggests that building CCSS reading assessments is more than a matter of selecting standards to test. Construct validity can best be achieved by building reading assessments in accordance with the following guidelines:

- **Test items should require analysis of text.** CCSS assumes students become proficient readers by reading texts. Consequently, assessment items should require reading and analysis of texts.
- **Favor holistic assessment of texts over selective measurement of standards.** Reading assessments should measure students’ ability to engage in close, sustained analysis of a text as a whole. However, this is difficult to achieve when selective measurement of standards governs the design of assessments.

For example, an assessment that measures students' ability to identify the main idea of a text multiple times will almost certainly include several texts; if those texts are only used to assess one standard, the resulting analysis is shallow and does not reflect the theory of reading advanced in the standards. Assessments should be designed so that students engage in complex, comprehensive analysis of fewer texts. In other words, prioritize holistic analysis of texts over selective measurement of standards when building assessments.

- **Include an appropriate balance of literary and informational texts.** Assessments should reflect the standards' recommended ratio of informational to literary texts. Include literary nonfiction on high school reading assessments.
- **Broaden the range of item types.** The assessments developed by PARCC and the SMARTER Balanced Assessment Consortium will include a variety of item types, such as selected-response, constructed-response, extended constructed-response, and performance-based tasks. Benchmark reading assessments should be designed with a similarly varied toolkit. Doing so broadens the range of standards that can be assessed. Moreover, some standards that can be assessed with selected-response items are better tested via other item types (e.g., a standard that calls for students to summarize a text is better assessed through constructed-response); assessment validity is increased when there is a logical connection between an element of the construct and the item type used to assess it (Haladyna, 2004).

V. Discussion

CCSS calls for all students to be prepared to be able to read at the college level upon graduation from high school. To accomplish this goal, teachers will need to shift their focus from shallow coverage of standards to sustained engagement with rich, complex texts.

This evolution of curriculum and pedagogy demands concurrent changes in assessment. Valid reading tests must invite students to read fewer, more challenging texts. Assessment items should help students proceed through comprehensive analysis of each text; moreover, the methods of assessment should follow logically from the skills and processes the standards expect students to demonstrate.

VI. References

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.

Council of Chief State School Officers (CCSSO) & National Governors Association (NGA). (2010). *Common Core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: Author.

Dunbar, S. B., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. *Applied Measurement in Education*, 4, 289—303.

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13—104). New York: American Council on Education and Macmillan.

Partnership for Assessment of Readiness for College and Careers. (2011, November). PARCC Model Content Frameworks, English Language Arts/Literacy, Grades 3—11. Retrieved from <http://www.parcconline.org/parcc-content-frameworks>