1

Language-Arts Standards and Testing: Get Specific!

E. D. Hirsch Jr.

Since Arkansas desires not only to improve its workforce but also to give all of its children a fair chance, no policy could be educationally more productive than a successful effort to raise test scores significantly in reading comprehension in the early grades. Much in educational progress depends on adequate reading comprehension test scores early on. Although some of my examples come from fourth grade, these remarks cover teaching and testing policy for the whole span of Arkansas instruction from Kindergarten through grade 12. Fourth grade happens to be the time when reading tests first begin to measure reading comprehension challengingly and reliably, which explains why reading comprehension scores of fourth graders are, on average, reliable predictors of a child's entire educational career.[1] Children's fourth-grade test scores in reading comprehension show a very high correlation with their later school grades, with their writing

1. Anne E. Cunningham and Keith E. Stanovich, "Early Reading Acquisition and Its Relation to Reading Experience and Ability 10 Years Later," *Developmental Psychology* 33, no. 6 (Nov 1997): 934–45.

ability, and finally, with their later income.[2] It has been definitively shown that a cumulative year-by-year approach to the knowledge children need is the single most effective method of raising achievement and narrowing the education gap between ethnic and racial groups.[3]

Tests drive schooling in the present day. The Arkansas *K–12 English Language Arts Curricular Framework* (rev. 2003) exercises its impact on Arkansas schools by supplying the criteria used by the makers of Arkansas' "criterion-referenced" tests—the technical term used to describe the *Benchmark* tests. Hence I will discuss the Arkansas *Framework* not just as a guide to classroom teaching but also as a determiner of what is on the *Benchmark* tests. These looming tests are the filters through which the *Framework* standards are viewed by teachers, students, and administrators. To consider "standards" without considering how they connect with the tests that measure them is an empty academic exercise. In the real world of schools, the tests are the de facto standards.

These recommendations to improve current Arkansas language-arts standards and tests are not intended as an implicit criticism of anyone, but rather of some outmoded ideas. The process-oriented ideas reflected in the Arkansas standards and tests are widely promulgated within education circles throughout the United States, but they are holding back progress in reading comprehension. By "a process-view of language-arts" I mean the kinds of formal skills set forth in the Arkansas *Framework*: such as the child will "use questions and monitoring to make meaning,"

2. Richard C. Hofstetter, Thomas G. Sticht, and Carolyn Huie Hofstetter, "Knowledge, Literacy and Power," *Communication Research* 26, no. 1 (Feb 1999): 58–80.

3. M. Duthoit, "L'enfant et l'ecole: Aspects synthetiques du suivi d'un echantillon de vingt mille eleves des ecoles," *Education et Formations*, no. 16 (1988): 3–13.

will "use inferences to make meaning," will "summarize and synthesize to make meaning." The assumption behind these guidelines is that reading comprehension is mainly a set of formal skills developed by practicing formal strategies such as "finding the main idea," "questioning the author," "clarifying," and "summarizing."

A scientific consensus in cognitive psychology now agrees that this process-view of literacy is altogether inadequate—a conclusion strongly supported by the low reading-proficiency rates in Arkansas and elsewhere that result from the process conception.[4] A recent analysis of the actual results of strategy instruction shows that a few days of teaching such skills yield the same small effects as months of such teaching do![5] The process conception

4. W. Kintsch, "The Role of Knowledge in Discourse Comprehension: A Construction Integration Model," *Psychological Review* 95 (1988): 163–182; Patricia A. Alexander, Jonna M. Kulikowich, and Sharon K. Schulze, "The Influence of Topic Knowledge, Domain Knowledge, and Interest on the Comprehension of Scientific Exposition," *Learning & Individual Differences* 6, no. 4 (1994): 379–97; Stepahnie Caillies, Guy Denhiere, and Walter Kintsch, "The Effect of Prior Knowledge on Understanding from Text: Evidence from Primed Recognition," *European Journal of Cognitive Psychology* 14, no. 2 (Apr 2002): 267–86; A. Garnham and J. Oakhill, "The Mental Models Theory of Language Comprehension," in *Models of Understanding Text*, ed. B. K. Britton and A. C. Graesser (Hillsdale, NJ: Erlbaum, 1996); Arthur C. Graesser and Rolf A. Zwaan, "Inference Generation and the Construction of Situation Models," in *Discourse Comprehension: Essays in Honor of Walter Kintsch*, eds. Charles A. Weaver III, et al. (1995): 117–39; Diana Christine Pulido, "The Impact of Topic Familiarity, L2 Reading Proficiency, and L2 Passage Sight Vocabulary on Incidental Vocabulary Gain Through Reading for Adult Learners of Spanish as a Foreign Language," abstract in *Dissertation Abstracts International* 61, no 10-A (May 2001): 3892; Donna R. Recht and Lauren Leslie, "Effect of Prior Knowledge on Good and Poor Readers' Memory of Text," *Journal of Educational Psychology* 80, no. 1 (Mar 1988): 16–20; and Rolf A. Zwaan and Gabriel A. Radvansky, "Situation Models in Language Comprehension and Memory," *Psychological Bulletin* 123, no. 2 (Mar 1998): 162–85.

5. Barak Rosenshine and Carla Meister, "Reciprocal Teaching: A Review of the Research," *Review of Educational Research* 64, no. 4 (Winter 1994): 479–530.

simply does not work. As a consequence, a very great deal of time in Arkansas language-arts classrooms is being spent unproductively. A significant change of policy is needed that places a greater emphasis on imparting the background knowledge which students need for proficient reading comprehension and effective writing. In brief I will try to lay out a system of standards and tests that will be simpler and cheaper to carry out than the present one, and will result in much higher reading comprehension scores.

There are two chief components of reading skill: decoding and comprehension. Proficiency in "decoding" means the ability to translate print into language accurately and fluently. Proficiency in "comprehension" means the ability, fluently and accurately, to understand what the language says. In the Arkansas *Framework*, these two elements are two strands under "Reading." They are labeled respectively "Foundations of Reading" and "Comprehension." I shall *not* discuss the "Foundations of Reading," the Arkansas state standards in the area of decoding, which are adequate. Schools are now well aware that the fastest and surest way to teach decoding is through systematic phonics. Most current textbooks are scientifically sound in teaching fluency of decoding, and much progress is being made in disseminating good instruction in K–2 in order to achieve proficiency in decoding. In this circumstance, the details of language in the "Foundations of Reading" section are not critical; what counts are the published, constantly improving decoding programs that teachers are actually using in classrooms.

Comprehension, though, is another matter. That is an area where students in Arkansas are not performing well. This may not be immediately apparent from reported scores. According to a recent RAND report comparing state scores, the state of Arkansas is currently reporting that 61 percent of its fourth graders read proficiently on the basis of the Arkansas *Benchmark* test. But that percentage contrasts strongly with the report of the

National Assessment of Education Progress (NAEP) which finds that only 28 percent of Arkansas fourth graders read proficiently.[6] This latter figure is the one that should be believed. The NAEP test results are trustworthy; the sampling procedures and the tests themselves are well validated and reliable. Policymakers should accept the NAEP conclusion that some 72 percent of Arkansas fourth graders now read below proficiency. If the recommendations of this chapter are put into effect, a very much larger proportion of them will read proficiently.

The Arkansas content standards for fourth-grade language-arts (and for all other grades) are remarkably lacking in content in the usual sense of the word. They describe processes only. These are the "content" criteria from the *Framework* on which the *Benchmark* test for fourth grade is based: "1. Students will comprehend, evaluate, and respond to works of literature and other kinds of writing which reflect their own cultures and developing viewpoints, as well as those of others. 2. Students will demonstrate a willingness to use reading to continue to learn, to communicate, and to solve problems independently." These vague, process standards give the teacher no real guidance, and insure that the actual subject matters learned by students from year to year will be uncertain and fragmented.

By specificity of content, I mean age-appropriate and time-appropriate treatment of topics as specific as the following: Grade 1: "Rumplestiltskin," Anansi Stories, "Jack and the Beanstalk," Aesop's *Fables*: "The Boy Who Cried Wolf," "The Fox and the Grapes," etc. Grade 2: "Beauty and the Beast," *A Christmas Carol*, Iktomi Stories, Paul Bunyan Stories, Greek Myths: Zeus, etc. Grade 3: "The People Who Could Fly," "Ali Baba and the Forty Thieves," Norse Myths: Odin, etc., Greek Myths: Jason,

6. Jennifer Sloan McCombs, et al., *Achieving State and National Literacy Goals: A Long Uphill Road*, A Report to Carnegie Corporation of New York (Santa Monica, CA: RAND Corporation, 2004).

Sword of Damocles, etc. Grade 4: *Robinson Crusoe*, Robin Hood, Legends of King Arthur; speeches: Patrick Henry, Sojorner Truth, etc. These content specifications should be integrated with equally specific topics in history and the arts and science. In fact, many nonfiction topics should be treated with language-arts, as the great reading researcher Jeanne Chall has recommended.

Currently the information to be tested on the Arkansas *Benchmark* tests concern abstract processes: "Students will use prior knowledge to extend reading ability and comprehension. Use specific strategies such as making comparisons, predicting outcomes, drawing conclusions, identifying the main ideas, and understanding cause and effect to comprehend a variety of literary genres from diverse cultures and time periods." The test itself consists of passages for the student to read followed by questions that probe the child's comprehension of the passages, by identifying the main idea, making comparisons, and so on.

Let me alert the readers of this short chapter to a key fact about this "criterion-referenced" reading test, and indeed about all reading tests. Although the creators of the *Benchmark* tests may have believed they were testing according to the stated criteria (i.e., the child's skills in finding main ideas, drawing conclusions, and making comparisons), such formal skill is not what is really being tested here. What is really being tested is the child's comprehension of the passage itself, and that comprehension chiefly depends upon the child having relevant knowledge. Assuming that the child has adequate decoding fluency, the strongest variable in reading comprehension is topic-familiarity. Children who accurately comprehend a passage on a reading test will accurately answer questions about it whether or not they have had training in formal strategies.

By the same token, children who read the passages inaccurately and with difficulty will not be able to answer questions well, no matter how much strategy training they have had. Assuming

adequate decoding skill, the chief determinant of a child's text comprehension is simply the child's familiarity with the general subject matter of the passage.[7] A child who is familiar with the subject matter and who understands the "situation model" that the passage represents will ace the test, whereas the child to whom the subject matter is unfamiliar will do poorly regardless of training in formal reading processes. To do reliably well in reading a diversity of texts, a child needs to possess broad general knowledge.

As a consequence of this fact about reading, the most valid tests of *general* reading comprehension ability, like the Iowa Test of Basic Skills (ITBS), and the Stanford 9 or 10, are tests which present a number of passages of different types on different subjects. The validity of a reading test is improved if it uses a variety of topics and text types, since such variety helps gauge accurately the child's general ability to read, while diminishing the effects of an accidental match or mismatch in familiarity between child and a single topic. The ITBS for grade four offers nine short passages, each followed by three to five multiple choice questions about the meaning of the passage. The Stanford 9 offers seven short passages each followed by five multiple choice questions. A child who knows something about a great many topics will therefore do well on a valid reading test, which gains its validity because it predicts how well a person can read a variety of texts.

General knowledge, not formal strategy, is the construct that is being measured by a valid reading comprehension test—how well the test taker can understand an unpredictable variety of grade-appropriate texts. A valid reading comprehension test is a general knowledge test.[8]

7. Patricia A. Alexander, Jonna M. Kulikowich, and Sharon K. Schulze, "The Influence of Topic Knowledge, Domain Knowledge, and Interest on the Comprehension of Scientific Exposition," *Learning & Individual Differences* 6, no. 4 (1994): 379–97.

8. John B. Carroll, "Psychometric Approaches to the Study of Language

It follows that the single most effective policy to improve scores on reading tests is to lead all students through a cumulative, knowledge-based curriculum that ensures their acquisition of general knowledge, especially the knowledge that is taken for granted in serious writing of all sorts.

Recommendations

1. The Arkansas *Framework* in language-arts should be revised to place much less emphasis on process and much more emphasis on imparting the general knowledge needed for reading comprehension.
2. The *Benchmark* tests by themselves or in combination with other tests like the Iowa Test of Basic Skills should measure not just general reading ability as indicated by multiple types of reading passages, but should also measure student progress in gaining the general knowledge needed for reading comprehension.
3. Both the criteria stated in the *Framework*, and the criterion-referenced tests that assess student mastery of them should *specify* grade-by-grade content in the ordinary meaning of the word "content"—that is, topics of knowledge like photosynthesis and Paul Revere's ride. Examples of states that specify grade-by-grade content include Massachusetts and Virginia, but, for historical reasons, *no* current state standards are good enough to serve as a model. Arkansas could be a leader in formulating truly specific and effective grade-by-grade content standards.
4. This specified content should be planned over all the grades

Abilities," in *Individual Differences in Language Abilities and Language Behavior*, eds. C. J. Fillmore, D. Kempler, and S. Y. Wang (New York: Academic Press, 1979).

so that there is a cumulative build-up of knowledge from kindergarten on, resulting in the student's gradual acquisition of the general knowledge that is most useful for general reading comprehension in the United States today. These content topics should be the criteria that are tested in the comprehension portions of *Benchmark* tests.

5. These grade-by-grade content-and-skills standards should be integrated in the various subjects, so they reinforce each other. The committees that decide on language-arts content should cooperate with the committees that decide content in history, science, and the arts.

Instituting these recommendations would elevate reading comprehension achievement not only beyond what Arkansas is currently achieving, but beyond what any other state is currently achieving, and will enable all schools to make adequate yearly progress for all groups under the terms of the No Child Left Behind Act. Many schools are not making adequate yearly progress in reading because school time is being used unproductively under the strategy theory of reading comprehension. Because children are gaining too little of the general knowledge they need for reading, they do not and cannot make adequate yearly progress.

In sum, if the *Framework* is revised to specify grade-by-grade content criteria, and if the *Benchmark* tests based on them should become truly criterion-referenced to these content criteria—in fact as well as in name—the result would be a significant rise in real-world reading and learning ability by Arkansas students.