

7-1-1990

Testing: We Need a New Perspective

W. John Harker
University of Victoria

Follow this and additional works at: https://scholarworks.wmich.edu/reading_horizons



Part of the Education Commons

Recommended Citation

Harker, W. J. (1990). Testing: We Need a New Perspective. *Reading Horizons: A Journal of Literacy and Language Arts*, 30 (4). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol30/iss4/8

This Article is brought to you for free and open access by the Special Education and Literacy Studies at ScholarWorks at WMU. It has been accepted for inclusion in Reading Horizons: A Journal of Literacy and Language Arts by an authorized editor of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



Testing Reading: We Need a New Perspective

W. John Harker

In their recent study of reading assessment, Farr and Carey (1986) observe that over the past several years "testing programs...have exploded on the educational scene" (p. 6). Those familiar with reading instruction and the assessment of children's reading development must agree. Testing programs at the district, state, and national levels have proliferated recently as more and more pressure is brought to bear on teachers to demonstrate in some quantifiable fashion their success (or lack of it) in teaching children to read.

An instance of this trend is *A Nation at Risk* (1983), published by the National Commission on Excellence in Education, where it is argued that "standardized achievement tests should be administered at major transition points from one level of schooling to another" (p. 18). Public and political receptivity to such arguments and the testing programs that follow from them can be seen, for example, in a bill passed by the Indiana Senate in 1984 which decreed that "student test scores would indicate a school by school ranking of Indiana's school corporation" (UPI, 1984). The frequency of use of standardized tests is indicated in a recent study by Carey (1985) which found that students going through the Rhode Island school system could normally expect to take between twelve and fifteen major test batteries during their school career. More generally, Anderson (1982) has estimated that students in American schools typically spend from two to six hours each year taking standardized tests. The *English*

language arts framework (California State Department of Education, 1987) contains the statement that "school districts may find useful the overview of students' skills and their use of language conventions provided by such objective instruments as criterion- and norm-referenced tests" (p. 36).

Teaching and testing

The frequency and growth in the use of standardized tests raises the question of their validity in measuring reading achievement. Put another way, do tests measure reading as we conceive it and teach it?

As educators know, our understanding of the reading process has undergone dramatic change over the past two decades. As teachers became familiar with the writings of Kenneth Goodman (1967) and Frank Smith (1971) in the late 1960's and early 1970's, their thinking about reading began to alter. There developed a growing recognition that reading does not involve the simple decoding of meaning represented in the text, but that it involves an interaction between background information the reader brings to the text and information the reader finds there. Through this interaction, the reader constructs meaning. This constructive emphasis has gained increasing support from research during the 1970's and 1980's (van Dijk, 1987; van Dijk and Kintsch, 1983; Just and Carpenter, 1980, 1987; Perfetti, 1985; Perfetti and Lesgold, 1977; Rumelhart, 1977) with the result that the traditional skills approach to teaching reading with its emphasis on basal readers and workbooks has been replaced by an emphasis on teaching specific comprehension strategies, replacing the limited content encountered in basal readers with literary selections and trade books, encouraging more independent reading, and combining reading and writing activities. In all of this, process has taken precedence over product.

But has it? Or has it when one examines standardized tests intended to determine reading achievement? The answer has to be "no." Despite the millions of dollars spent annually on tests, their pervasive use in determining student achievement, their influence on curriculum planning, and the enormous public and media attention given to the outcomes of testing, standardized reading tests remain locked in a concept of reading which does not coincide with current knowledge of the reading process.

Perhaps the most obvious example of this lies in their skills emphasis. Rather than assessing the underlying processes through which readers construct meaning, the majority of standardized tests measure children's performance of various arbitrarily prescribed reading skills which research has shown to have little or nothing to do with reading or learning to read. The result is that these tests measure an artificially fragmented and contrived construct of the reading process rather than the highly integrated interactive one which research repeatedly reveals reading to be.

Another way standardized tests differ from our current understanding of the reading process is by trying to eliminate the effect of background information. As Johnson (1983, 1984) has shown, test makers do this by including a broad range of topics in test passages, by eliminating questions which they think students with greater background information can answer more easily, and by using statistical models based on estimates of subgroups' knowledge of the topic included. And yet, in doing so, test makers attempt to eliminate one of the most important elements in reading comprehension. Although background information clearly influences test performance, it does so because it is a fundamental component of the reading process. The removal of its influence is therefore impossible in the valid measurement of reading.

A related shortcoming of standardized tests is the contrived nature of the reading passages which they employ. Rather than relying on excerpts from authentic language sources such as children's literature and nonfiction, test makers often use reading passages which are specifically designed for their tests. However, as Nystrand (1987) has shown, these passages are frequently unrealistic and trivial in their content. Moreover, as revealed in the research of Fillmore and Kay (1983) and Langer (1987), these passages are often puzzling, inconsistent, and conflicting in the information they contain. The result is that the reader is misled, forced to second-guess meaning, and to adopt separate test-taking strategies which are unlike those employed in normal reading.

Another difficulty with standardized tests is their insensitivity to inferential understanding. This shortcoming has long been recognized, even before constructivist models of reading evolved. However, with the evolution of these models and the research which supports them, the emphasis of standardized tests on the measurement of literal comprehension over inferential comprehension is an even more serious shortcoming. Research into the nature of reading and learning to read has repeatedly shown the importance of inference in constructing meaning (Dewitz, Carr and Patberg, 1987; Hansen, 1981).

Problems

The major danger in the use of standardized tests which vary so markedly from what we know about the reading process is that they limit instruction rather than further it. This limiting influence shows itself in several ways. One of these is through the almost ritualistic fashion in which standardized tests are administered, and the manner in which their results

are received and interpreted. It is not uncommon to witness the administration of a standardized test when the purpose for testing has never been clearly established and the relationship of the particular test used to the instructional program has never been considered. And, often, when the results come in, they are accepted as truth, as a commendation or condemnation of the instructional program regardless of the validity of the test for evaluating that program. Such testing wastes time of teachers and students alike since it reveals little or nothing about reading achievement in the particular educational setting in which it is used.

Another problem comes from the reaction to test scores by teachers who are unaware of the tests' shortcomings or who, because of administrative and public pressures, feel inhibited from challenging their validity. These teachers teach what tests measure without regard for the incompatibility between what they actually measure and current knowledge of the reading process. They remain bound by a skills approach which does little more than prepare children for success on subsequent reading tests. An associated problem lies in the way test scores are interpreted. Students who have been taught to use the contrived skills set by standardized tests may well achieve higher test scores than those who have been taught constructive reading strategies. The result is to discredit the teaching of these strategies in the eyes of those for whom test scores are the beginning and end of reading instruction, and to further entrench instruction in the meaningless reading skills which tests measure.

A further difficulty with standardized testing is that the range of reading skills measured by any single test is significantly less than the range of skills taught through the traditional skills-building basal program. This has been a criticism

of standardized tests for decades. However, it becomes even more telling in light of what we now know about the reading process and the fallacy of measuring specific skills in the first place. The question which faces us now is, given the complexity of the reading process as revealed through recent research and theory, can any one standardized test or even battery of tests provide an adequate total picture of children's reading behavior?

Another limitation of standardized tests is that, due to their perceived authority, they diminish teachers' confidence in using informal tests they make themselves. However, teacher-made tests are often better at revealing children's reading performance in the context of the particular tasks demanded of them in normal classroom learning situations. The artificiality of standardized test administration — the tension created by the unusual situation of test administration with its strange-looking booklets, the pressure of timing, the stilted instructions, the unnatural content of the reading passages, and so on — reduces the validity of these tests. Informal tests overcome much of this artificiality through the natural and informal manner in which they are administered, and the similarity between their content and the reading material children normally encounter in the classroom. Through informal testing teachers are able to integrate the process of instruction with the process of evaluation so that the two become almost indistinguishable. But pressure for the formal quantification of reading performance through the authority of standardized tests often leads teachers to defer to these tests and diminish the value of their own tests.

Solutions

In the face of these problems, it is not surprising that informed teachers have become increasingly disillusioned with the use of standardized tests in their classrooms. And yet it

seems clear that testing in some form is here to stay. Not only is testing frequently represented as the only sure way to guarantee quality in education, but this belief is at least tacitly encouraged by test publishers for whom standardized tests are profitable big business.

Given this situation, the challenge facing reading teachers is not to eliminate testing, but to make it more responsive to valid educational goals as represented in our current understanding of how children read and learn to read. There are many ways this can be done, and although none of them is easy, what follows are some suggestions.

First, teachers must become knowledgeable about current concepts of the reading process in order to be articulate and informed agents of change in testing. Little will be gained by advocating change when teachers themselves are not clear as to what the nature of this change should be, in terms of what is currently known about the reading process. Therefore, self-education is a necessary first step toward strong and informed advocacy.

Once a sound knowledge base for advocacy is established, teachers should not remain confined to an audience of their peers in expressing dissatisfaction with standardized testing. Support for standardized testing most often comes from uninformed public opinion, and teachers' opposition to this testing must therefore be voiced in the public arena. Speaking to service clubs, church groups, parents' organizations, and similar audiences is a role teachers should actively assume if they expect their side of the testing argument to be heard. What all this means is that teachers must become more vocal advocates of valid testing. Too often teachers voice their concerns among themselves without "going public."

Within their professional activities, teachers can resist the tendency for testing to determine the reading curriculum. When testing establishes the ends and means of instruction, when what tests measure constitutes the goals of instruction, and when deficiencies revealed by tests determine instructional objectives, the process is circular. This circle is even more insidious when what standardized reading tests actually measure under the designation "reading" is in fact something quite different from what current research and theory reveals reading to be. The fundamental issue here is whether standardized tests, or teachers, should determine educational outcomes and educational practice.

In recognition of the fact that to ignore the call for testing of some type is unrealistic in today's educational climate, teachers should insist on the validity of their own informal tests and the information these tests provide. Informal tests can combine process and product information to a far greater extent than standardized tests. These tests can be designed so that the particular reading strategies demanded of children in specific learning situations can be observed, and the outcome of this learning can be determined. Gone is the artificiality of standardized tests and their distance from the normal instructional procedures of the classroom. Rather, what can be called "situational validity" is established as children work in normal learning situations performing test tasks in such a way that they are often not even aware they are being evaluated. Yet, as a result of such testing, teachers gain insights into the process through which children are learning as well as the product of this learning. Suggested formats for the development of informal tests are available from numerous sources including Ahrendt and Haselton (1973), Flint-Ferguson and Youga (1987), Royer, Greene, and Sinatra (1987), Simpson (1987), Voix (1968), and Wood (1985).

Closely related to the use of informal tests is a needed emphasis on the teacher as observer. Observing children's reading behavior, their ease and attention when performing reading tasks, the kinds of material they like to read, the manner in which they communicate their feelings about what they have read, and the choices they make in response to new reading material of varied difficulty and content — all provide the observant teacher with a wealth of information about children's reading. This information, combined with information derived from teacher-made tests, can serve as the basis for qualitatively rich reports of children's reading progress which can be made to answer even the most persistent demands for testing.

In all of this, a new perspective on testing is called for. The responsibility for bringing about this perspective lies primarily with teachers, with a clear recognition of their central role in evaluating children's reading, and an understanding of the necessity to test reading in ways which are consistent with what we know about how children read and learn to read.

References

- Ahrendt, K. M. & Hasleton, S. S. (1973). Informal skills assessment for individualized instruction. *Journal of Reading*, 17, 52-57.
- Anderson, B. (1982). Test use today in elementary and secondary schools. In A.K. Wigdor & W.R. Barner (Eds.), *Ability testing: Uses, consequences, and controversies, part 2*. Washington DC: National Academy.
- Carey, R. F. (1985). *Program evaluation as ethnographic research*. Providence RI: Rhode Island Department of Education.
- Dewitz, P., Carr, E. M., & Patberg, J. P. (1987) Effects of inference training on comprehension and comprehension monitoring. *Reading Research Quarterly*, 22, 99-122.
- California State Department of Education. (1987). *English language arts framework*. Sacramento CA.
- Farr, R., & Carey, R. F. (1986). *Reading: What can be measured?* Newark DE: International Reading Association.
- Fillmore, C.W., & Kay, P. (1983). *Test semantic analyses of reading comprehension tests*. Berkeley CA: Institute of Human Learning, University of California at Berkeley.

- Flint-Ferguson, J., & Youga, J. (1987). Making evaluation a part of the learning process. *Journal of Reading*, 31, 140-145.
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6, 126-135.
- Hansen, J. (1981). The effects of inference training and practice on young children's comprehension. *Reading Research Quarterly*, 16, 391-417.
- Johnson, P. (1983). *Reading comprehension assessment: A cognitive basis*. Newark DE: International Reading Association.
- Johnson, P. (1984). Prior knowledge and reading comprehension test bias. *Reading Research Quarterly*, 19, 219-239.
- Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston MA: Allyn and Bacon.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Langer, J. A. (1987). The construction of meaning and the assessment of comprehension: An analysis of reader performance on standardized test items. In R. O. Freedle & R. P. Duran (Eds.), *Cognitive and linguistic analyses of test performance*. Norwood NJ: Ablex.
- The National Commission on Excellence in Education. (1983). *A nation at risk*. Washington DC: U.S. Department of Education.
- Nystrand, M. (January, 1987). Review of selected national tests of writing and reading. Paper prepared at the National Center on Effective Secondary Schools, School of Education, University of Wisconsin-Madison.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A., & Lesgold, A. M. (1977). In M. A. Just & P. A. Carpenter (Eds.), *Discourse comprehension and sources of individual differences*. Hillsdale NJ: Erlbaum.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale NJ: Erlbaum.
- Royer, J. M., Greene, B. A., & Sinatra, G. M. (1987). The sentence verification technique: A practical procedure for testing comprehension. *Journal of Reading*, 30, 414-422.
- Simpson, M. L. (1987). Alternative formats for evaluating content area vocabulary understanding. *Journal of Reading*, 31, 20-27.
- Smith, F. (1971). *Understanding reading*. New York: Holt, Rinehart, and Winston.
- United Press International. (February 14, 1984). Bill would rank schools by test scores. *Herald Telephone*, 12. Bloomington IN.
- van Dijk, T. A. (1987). Episodic models in discourse processing. In R. Horowitz and S. J. Samuels (Eds.), *Comprehending oral and written language*. New York: Academic Press.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discoursed comprehension*. New York: Academic Press.
- Voix, R. G. (1968). *Evaluating reading and study skills in the secondary classroom*. Newark DE: International Reading Association.

Wood, K. D. (1985). Free associational assessment: An alternative to traditional testing. *Journal of Reading*, 29, 106-111.

W. John Harker is a member of the Faculty of Education at the University of Victoria in Victoria, British Columbia, Canada.

IRA DELEGATES SPEAK OUT ON ASSESSMENT

At the 1990 conference of the International Reading Association, held in May in Atlanta Georgia, the 389-member delegate assembly voted unanimously to oppose "the proliferation of school-by-school, district-by-district, state-by-state, and province-by-province comparison assessments," noting specifically the biennial assessments by the National Assessment of Educational Progress (NAEP), and congressional mandates for comparison studies of NAEP test results which will further increase the cost of the estimated \$19 million allotted for the 1992 assessment.

Judith Thelen, of Frostburg State University in Maryland, who is IRA president-elect, stated, "Reading educators are not opposed to measuring progress. But outmoded tests are not testing what we are now teaching." Current IRA president, Carl Braun, of the University of Calgary in Alberta Canada, asserted, "External control over assessment, especially inappropriate use of large scale assessment data, is recognized as a threat to the work of teachers and ultimately the welfare of our children. This action by our delegates attests to the determination of our members to stand firm on issues that directly impact the lives of teachers and children."

Probably the most incisive comment on the current emphasis on mandated, extensive, continuing testing came from Heather Fehring, IRA delegate from Australia: *"As any wise old farmer will tell you, you don't fatten your lambs simply by weighing them."*
