What Is Evidence-Based Reading Instruction?

This brochure may be purchased from the International Reading Association in bulk quantities, prepaid only. (Please contact the Association for pricing information.) Single copies are free upon request by sending a self-addressed, stamped envelope. Requests from outside the U.S. should include an envelope, but postage is not required. Single copies also can be downloaded free for personal use through the Association’s website at www.reading.org (requires Adobe’s Acrobat Reader).

© 2002 International Reading Association
Brochure design by Linda Steere
Cover photo by EyeWire

A POSITION STATEMENT OF THE INTERNATIONAL READING ASSOCIATION

“Evidence-based reading instruction” is an objective, systematic, peer-reviewed practice.
There are few instructional tasks more important than teaching children to read. The consequences of low achievement in reading are costly both to individuals and society. Low achievement in literacy correlates with high rates of school dropout, poverty, and underemployment (Snow, Burns, & Griffin, 1998; Wagner, 2000). The far-reaching effects of literacy achievement have heightened the interest of educators and noneducators alike in the teaching of reading. Policymakers, parents, administrators, and teachers seek the same end—to provide literacy instruction that is most likely to lead to high rates of achievement for all children.

As we pursue this goal, we must be mindful of the critical lesson provided by investigations of the past and of the present: There is no single instructional program or method that is effective in teaching all children to read. Rather, successful efforts to improve reading achievement emphasize identification and implementation of evidence-based practices that promote high rates of achievement when used in classrooms by teachers with diverse instructional styles with children who have diverse instructional needs and interests (Bond & Dykstra, 1967/1997; National Clearinghouse for Comprehensive School Reform, 2001).

Also, as we seek effective programs and practices, we must remain mindful of the powerful influence teachers have. Time and again, research has confirmed that regardless of the quality of a program, resource, or strategy, it is the teacher and learning situation that make the difference (Bond & Dykstra, 1967/1997). This evidence underscores the need to join practices grounded in sound and rigorous research with well-prepared and skillful teachers.

**What does the term evidence-based reading instruction mean?**

In its simplest form, evidence-based reading instruction means that a particular program or collection of instructional practices has a record of success. That is, there is reliable, trustworthy, and valid evidence to suggest that when the program is used with a particular group of children, the children can be expected to make adequate gains in reading achievement. Other terms that are sometimes used to convey the same idea are research-based instruction and scientifically based research.

This relatively simple concept becomes more complicated when we attempt to define the types of evidence that are reliable and trustworthy indicators of effectiveness. The central question is, What counts as evidence of success? In general, educators agree that such evidence should be as follows:

- **objective**—data that any evaluator would identify and interpret similarly
- **valid**—data that adequately represent the tasks that children need to accomplish to be successful readers
- **reliable**—data will remain essentially unchanged if collected on a different day or by a different person
- **systematic**—data that were collected according to a rigorous design of either experimentation or observation
- **refereed**—data that have been approved for publication by a panel of independent reviewers

In addition to evaluating the quality of the data by which programs or practices are judged, teachers also must examine the generalizability, or fit, of the evidence. In other words, teachers might ask if the children in their classrooms closely resemble the children from whom the evidence was collected: Are they the same age? Do they have similar language and cultural backgrounds? Do they have similar learning profiles? Teachers might also ask if the learning contexts are the same: Are class sizes and teacher–student ratios similar? Is the allocation of instructional time and resources similar? Do teachers have similar funds of knowledge? Has more than one study produced particular findings? If the answer to all of these questions is yes, then teachers might conclude that there is a good fit and that their students might be expected to make similar achievement gains with the same program or practice. If, however, the answers to some or all of these questions is no, then it is difficult to predict whether similar results might be achieved.

Research studies used to collect evidence about programs and practices may have a variety of designs. In general, studies that demonstrate effectiveness using experimental designs (studies that compare results from the program or practices of interest to results from a control group with random assignment to the groups) and quasi-experimental designs (studies that do not use random assignment to the program or comparison group, but use adequate statistical procedures to control preexisting differences) give the strongest evidence of effects of a program or practice on the “average” student—particularly when the studies are carried out in naturalistic environments. Quantitative studies such as these generally investigate program effects on relatively large numbers of students. In addition, they can be aggregated by using meta-analysis. In contrast, qualitative studies typically focus on small samples or on individuals and are especially valuable in helping teachers understand how particular programs or approaches affect individuals who may not represent the mainstream or average student.

However, no single study ever establishes a program or practice as effective; moreover, it is the convergence of evidence from a variety of study designs that is ultimately scientifically convincing. When evaluating studies and claims of evidence, educators must not determine whether the study is quantitative or qualitative in nature, but rather if the study meets the standards of scientific research. That is, does it involve “rigorous and systematic empirical inquiry that is data-based” (Bogdan & Biklen, 1992, p. 43)?
What is the difference between evidence-based programs and evidence-based practices?

The quest to find the “best practices” for teaching reading has a long and quite unsuccessful history. Most notable among such efforts is a group of studies conducted in the mid-1970s that became known as the New York City Reading Project (Bond & Dykstra, 1967-1997). This series of U.S. federally funded investigations examined popular approaches to teaching beginning reading. Included were examinations of basal reading, phonics, language experience, and linguistics approaches to reading instruction. The collection of 27 studies comparing different methods and materials found as many differences between and among teachers using the same method as there were between and among teachers using different programs or approaches, leaving authors unable to identify a “best” program. Instead, the results led the authors to conclude, “Children could be taught by a variety of methods and approaches...No one approach is so distinctly better in all situations and respects that it should be considered the one best method and the one to be used exclusively.” (Bond & Dykstra, 1967).

Indeed, many large studies have come to similar conclusions. For example, consider the recent findings related to the evaluations of Comprehensive School Reform. Once again the focus was on reading programs and methods, and the findings echo those of the First-Grade Studies, that “no models had uniformly positive effects, and no models had uniformly negative or neutral effects. In other words, no model worked in every case and every situation” (National Clearinghouse for Comprehensive School Reform, 2001, p. 21).

Despite many attempts at program studies in the years since the First-Grade Studies, and many claims of program excellence, literacy scholars (e.g., Allington, 2001; Stahl, Duffy, Nelson, & Roehler, 1998) argue that careful examination and systematic analysis of such studies reveals the use of either flawed designs or selective reporting of the available data. Furthermore, attempts to find the “right program” for large-scale implementation is complicated by the diversity of student needs, teaching styles, and classroom conditions that exist in any school or group of schools.

Programs referred to as “best programs” have centered largely on the notion that teachers use, attempts to identify best practices have focused on the actions teachers take and the practices in which they routinely engage students. In contrast to the disparate findings of studies designed to identify best programs, examinations of best practices have led to highly consistent results when such studies have been rigorously designed and systematically analyzed and compared. The results of the First-Grade Studies again provide a relevant starting place. Although findings failed to show superiority of any particular program or approach, evidence did indicate strong relationships between particular practices and high achievement. Most recently, the National Reading Panel (National Institute of Child Health and Human Development, 2000) took a similar approach to its study of effective instructional reading, examining evidence related to practices in phonemic awareness, phonics, fluency, vocabulary, and comprehension instruction. They found 22 phonics programs that varied along several dimensions that were effective. The results support a conclusion that it is particular practices and not the specific programs that are effective.

Current critical and comprehensive research reviews (e.g., Gambrell & Mazzoni, 1999; Guthrie & Alvermann, 1999; Kamil, Mosenthal, Pearson, & Barr, 2000; National Institute of Child Health and Human Development, 2000) took a similar approach to its study of effective instructional reading, examining evidence related to practices in phonemic awareness, phonics, fluency, vocabulary, and comprehension instruction. They found 22 phonics programs that varied along several dimensions that were effective. The results support a conclusion that it is particular practices and not the specific programs that are effective.

Current critical and comprehensive research reviews (e.g., Gambrell & Mazzoni, 1999; Guthrie & Alvermann, 1999; Kamil, Mosenthal, Pearson, & Barr, 2000; National Institute of Child Health and Human Development, 2000) took a similar approach to its study of effective instructional reading, examining evidence related to practices in phonemic awareness, phonics, fluency, vocabulary, and comprehension instruction. They found 22 phonics programs that varied along several dimensions that were effective. The results support a conclusion that it is particular practices and not the specific programs that are effective.

What resources might be useful when examining evidence to support particular programs or practices?

A list such as the one presented above provides an important starting point in the development of evidence-based reading instruction. But how might we learn more about each of these programs and the steps toward effective implementation? Rigorous, peer-reviewed, comprehensive research syntheses provide an excellent starting place for teachers, administrators, and policymakers who wish to learn more about effective teaching of reading. Such synthesized are important and useful because they are based on comprehensive and systematic reviews of many studies, and allow us to predict which practices and programs are used under similar conditions with children similar to those who participated in the reported investigations.

There are at least three types of research syntheses: large-scale reviews, small-scale reviews conducted by a single research team appointed by a funding agency; edited handbooks, generally compiled by a team of researchers who invite professional colleagues to provide comprehensive reviews of particular topics within a series of chapters, and topic-oriented reviews. Individual analyses may be published as book-length monographs, as articles in refereed research journals, or as chapters in edited volumes. The following list gives examples of these types of works, as well as titles in refereed research journals.

Large-scale, U.S. federally funded research reviews


Edited handbooks


References


