

10 THINGS EVERY LITERACY EDUCATOR SHOULD KNOW ABOUT RESEARCH

Nell K. Duke ■ Nicole M. Martin

“Research-based,” “research-proven,” “scientifically based”—in the reading world these days, it seems that the term *research* is being used

everywhere. It is also being misused and misunderstood. In fact, we are encountering a growing number of literacy educators who are dismissing research altogether, based on the belief that research is simply a propaganda tool for those trying to push a particular approach to reading and writing instruction. The purpose of this article is to argue for the value of research for literacy educators, including classroom teachers, coaches, specialists, and professors, and provide some information that may help us make better use of research and, at the same time, guard against misuse as we plan for and teach literacy.

Specifically, we discuss 10 things that we believe every literacy educator should know about research:

1. What research can do.
2. What research is.
3. What research is not.
4. The difference between *research-based* and *research-tested*.



5. Many kinds of research have valuable contributions to make to our understanding of literacy learning, development, and education.

6. Different kinds of research are good for different questions.

7. High-quality research has a logic of inquiry.

8. Conclusions drawn from research are only as sound as the research itself.

9. Where and how research is published or presented requires particular attention.

10. Educational research proceeds through the slow accumulation of knowledge.

1. What Research Can Do

Given all there is to know about research, and all the dangers of misrepresentation and misuse of research, you might wonder whether research is worth paying

Nell K. Duke is a professor of teacher education and educational psychology and codirector of the Literacy Achievement Research Center at Michigan State University, East Lansing, USA; e-mail nkduke@msu.edu.

Nicole M. Martin is an assistant professor of reading education at the University of North Carolina at Greensboro, USA; e-mail nmartin779@gmail.com.

“We believe that research should be seen as an essential guide to policy and practice.”

attention to at all. Educators frequently tell us that they just go with what works for them, that they trust what their mentor teacher does over what some research study says, or that research does not really apply to them or their setting. Although we certainly recognize the value of experience and that research will never provide all the solutions or answers, we believe that research should be seen as an essential guide to policy and practice. Concentrating here on practice, the following are some of the reasons for this belief.

Our Experiences Alone May Misguide Us

For a very long time, people believed that Earth was flat, and that was (and still is) a reasonable conclusion to draw from personal, individual observations. Research, however, eventually proved this view to be incorrect. In reading education, we have these kinds of examples as well, although perhaps not as dramatic. For instance, for a long time, we thought that persistent word reading difficulty, often called dyslexia, was primarily a visual problem. To illustrate, we thought that reading *star* as *rats* reflected a problem with the visual processing of print. We now understand that most word reading difficulties are actually caused by problems in phonological or auditory processing rather than visual processing (Snowling, 2000).

Similarly, we frequently visit classrooms in which teachers assign students to look up a list of vocabulary words in the dictionary and write the definition and a sentence containing the word.

This is a very commonsense practice. If we asked people on the street to suggest a good way to teach children new vocabulary, many would no doubt recommend exactly this practice. However, as it turns out, this practice often proves to be inferior to other approaches that although less common and arguably less commonsensible, actually result in greater vocabulary learning (e.g., Bos & Anders, 1990). Just as research has limitations, so too does common sense and individual judgment.

Sometimes We Do Not Know What We Do Not Know

We may not recognize experientially that something is a problem. A practice may be so widespread that it has been accepted as conventional wisdom, whether or not it is effective or true. For example, for many years in the United States, reading materials used in the primary grades were overwhelmingly stories; informational text, among other genres, was strongly neglected (Duke, Bennett-Armistead, & Roberts, 2003). Research is an important tool for documenting phenomena such as this (e.g., Duke, 2000; B. Moss & Newton, 2002).

Indeed, many people have shared with us that they had not really thought about how little experience children were getting within informational text in the primary grades. Others had thought about this but operated with assumptions that young children should read stories first and that reading stories was more “natural” for children than reading other genres. It turns out that this is not the case (e.g., Duke &

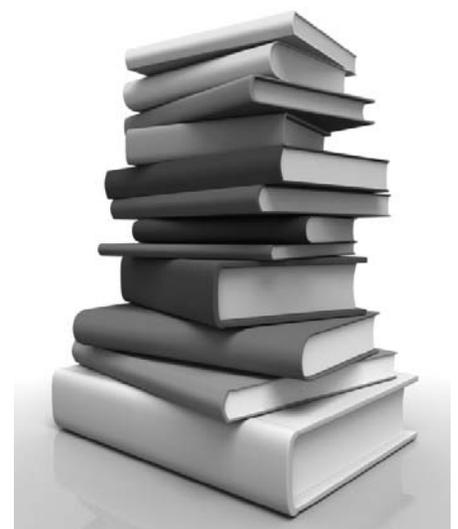
Kays, 1998; Pappas, 1993), and research was again an important tool for examining an assumption that was taken for granted.

Research Allows Us to Take a Longer Term View Than Our Personal Experiences May Allow

A teacher typically only has one to two years with a student, and it is difficult to systematically monitor a student’s progress after that time. This means that the teacher cannot observe the long-term outcomes of his or her practices. Research can do this. Researchers can track students as they move through the grades, for example, by administering the same assessments over multiple timepoints to understand long-term growth.

Research Allows Us Into Places and Situations That We May Not Be Able to Observe Otherwise

The constraints of daily life mean that many literacy educators simply cannot engage in practices such as spending long periods of time in a child’s home

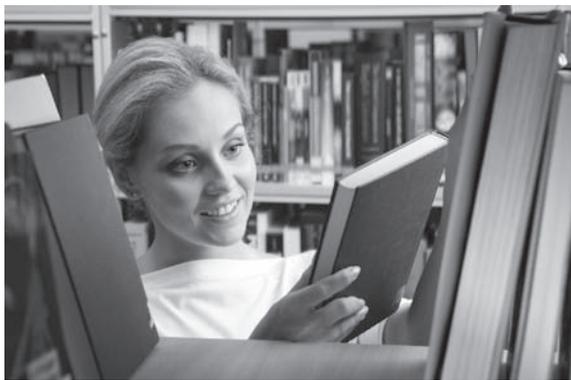


or observing what students do and do not do when visiting their local library. Research can provide some of this information for us. For example, Perry (e.g., 2009) spent 18 months in the homes, churches, and other settings of Sudanese refugee families. Her research provided valuable insights about the literacy practices in which these children were and were not engaged and ways in which these practices were and were not consistent with school literacy practices. This work would have been nearly impossible for a full-time literacy educator to conduct.

Research Allows Us to Pool Our Numbers and Experiences

As individuals, we simply may not have enough experience with a question to feel confident that we have a reasonable answer. For example, a classroom teacher may not have had enough students with disorders on the autism spectrum to feel confident about which instructional approaches might be most efficacious, or a literacy coach may not have had enough experience working with teachers to know what forms of professional development are likely to result in the greatest change in teacher practice. Research allows us to pool data across many sites and settings to address important questions about phenomena and practices.

In the end, we believe that research has value. It is worth learning about. It is worth paying attention to, even when it requires sorting out misrepresentation



and misuse. The remainder of this article focuses on knowledge that we believe can help literacy educators make use of this valuable tool.

2. What Research Is

Put simply, in our view, research is the systematic collection and analysis of data to address a question. Collecting and analyzing pre- and posttest data in multiple classrooms to find out which approach to writing instruction was more effective, collecting and analyzing data on the instructional practices of highly effective reading teachers, and collecting and analyzing data to learn about how a particular family engages in literacy practices with their children—all of these, assuming they are done systematically, are examples of research.

The ultimate purpose of literacy research is to deepen understanding of and thus improve literacy education. Of course, this is the ultimate purpose of literacy educators as well. In fact, at their essence, the activities in which literacy educators and literacy researchers engage are often not that different:

intensive, long-term observation of a student or students; close observation of student learning following implementation of a new instructional approach or materials; careful examination of the reading scores of students in a school district over the preceding years; thoughtful analysis of student talk during class activities; and so forth. Both literacy educators and literacy researchers engage in these activities with the end goal of improving literacy teaching and learning. We are, as the expression goes, on the same team.

Despite being on the same team, teachers and researchers do not collaborate or communicate as much as we might like. Teachers often tell us that they find researchers intimidating, and some researchers find great teachers intimidating! However, we have observed that some of the most compelling research arises from collaborations between researchers and teachers, when teachers and researchers share insights and burning questions that they have about practice and perhaps tinker together toward answers. For example, the famous research on concept-oriented reading instruction, or CORI, began as a collaboration between researchers at the University of Maryland and teachers in public elementary schools in Prince George's County, Maryland (e.g., Guthrie, Wigfield, & Perencevich, 2004).

You cannot tell whether something is research from where it is published. Periodically, research journals, such as *Reading Research Quarterly*, include pieces that are not research, such as essays on possible directions for future research in a particular area. Similarly, some articles in a practitioner journal, such as *The Reading Teacher*, may report on a research study with the intent of

“Both literacy educators and literacy researchers engage in these activities with the end goal of improving literacy teaching and learning.”

bringing findings and implications of that study to a broader audience. That said, there are a set of elements that you should expect to find, in one form or another, in any report of research, although they may not be neatly labeled as such:

- A statement of the research question(s) and/or purpose(s)
- A rationale for the study (i.e., information about why the question was asked, the problem the research was meant to investigate or address, and the research and theory that came beforehand)
- A description of the methods used to collect data to address the question(s)
- A description of the methods used to analyze the data collected to address the question(s)
- Results of these analyses
- Conclusions that the researchers have drawn based on the results
- Implications of those conclusions for practice
- Limitations of the study
- Directions for future research

To help identify these different components of a report of research, Figure 1 provides a note-taking sheet. If a piece of writing or a presentation does not include most of these elements—for example, if it appears to be exclusively one person’s opinion about effective practices or a memoir of his or her teaching experience—then this is a good indication that it probably is not research. This does not mean that it is not valuable, just that it is not research.

There are many kinds of research (see Table 1). For example, some research is focused on examining a single case, such as a single student or classroom. That

research is not intended to generalize or apply to all students or classrooms; its contributions lie in other areas, such as building theory. Other research, such as some survey and assessment research, is conducted with nationally or internationally representative samples and intended to have findings that generalize or apply to an entire population. No one kind of research is *less* than research of another kind. Different kinds of research address different kinds of questions and provide different kinds of information and insights. This variety allows us to gain deeper insights into a particular topic and begin to solve complex problems. We discuss this idea in greater depth in item 5.

You may notice that the term *teacher research* does not appear in Table 1. This is because the term refers to who is doing the research rather than the kind of research being done. Teachers, like researchers, can do any of the kinds of research listed in the table, although some kinds are undoubtedly more practical than others. Teachers’ goal for the research may be more local—generally, although not always, to inform their own practice but not to generalize to the practice of others—but their methods may be the same as those of researchers who hope to impact practice more widely.

Finally, we should note that some people use the terms *scientific research* or *scientifically based research*. When probing, we find that different people intend quite different meanings by these terms. In fact, we urge educators who hear others use these terms to ask what is meant by them. In our view, these terms are redundant. It is not as though, as some seem to believe, some kinds of research are scientific, and some are not. By (our) definition, *all* research is scientific.

3. What Research Is Not

Recognizing what research is *not* is as important as recognizing what research is. One common misconception that we encounter is the notion that if something is written by a researcher, it automatically counts as research. If a novelist writes a shopping list, its authorship does not make it a novel. If a researcher writes an article, book, or chapter, this fact alone does not make it research. Similarly, just because a researcher is presenting a set of practices does not mean that those practices are research-tested or research-based.

Also, just because a researcher develops an approach, works on an educational product, or endorses an instructional technique does not mean that the result is research-tested or even research-based. A researcher may be trying something completely new, may have only very limited influence on a product, or may be making a less-than-informed guess about the effectiveness of an instructional technique. Not everything a researcher touches becomes research or is even research-based, so we need to interrogate writings or products with a researcher’s imprimatur just as we would anyone else’s.

A second misconception that we have encountered is the notion that if something is written or developed by someone who has a position at a university, then the work is necessarily research or research-based. For example, we recently heard a teacher claim that the effectiveness of a particular approach to writing instruction must be research-proven because it was developed by a person who works at a university. In reality, there are many different kinds of positions at universities and colleges. Many of these do not require position

10 THINGS EVERY LITERACY EDUCATOR SHOULD KNOW ABOUT RESEARCH

Figure 1 Research Study Summary Sheet

Complete reference for the study _____

Purpose and/or questions	
Methodology	
Participants	
Context	
Data collection procedures	
Measures (if applicable)	
Materials (if applicable)	
Intervention (if applicable)	
Data analysis procedures	
Findings	
Limitations	
Implications for policy, practice, research, and other	
Other information for evaluation of study quality	

10 THINGS EVERY LITERACY EDUCATOR SHOULD KNOW ABOUT RESEARCH

Table 1 Brief Descriptions and Standards of Quality for Some Common Research Methodologies

Methodology	Description	Five quality indicators
Assessment development	These studies examine the reliability and validity of assessments, attitude surveys, and other research tools. Along with other activities, researchers typically give the assessment to a number of participants and then perform statistical analyses to examine its validity and reliability.	<ol style="list-style-type: none"> 1. The targeted concepts have been carefully defined. 2. Evidence has been provided suggesting that current research/theory has thoughtfully informed the assessment's creation, and other people have reviewed the assessment. 3. Multiple types of reliability and validity have been discussed. 4. The assessment includes uncomplicated and unbiased questions. Enough, but not too many, questions have been included to understand the targeted concepts. 5. The assessment includes instructions that are easy to understand and follow.
Case study	Case studies seek to describe naturally occurring phenomena. These studies often focus on a single or small number of cases, such as one classroom or three reading groups at one grade level. Researchers typically identify themes or patterns, rather than making claims about cause–effect relationships.	<ol style="list-style-type: none"> 1. It is easy to envision the context in which the research has occurred. Careful, vivid, and thorough descriptions have been included. 2. Potential ethical issues, such as the researchers' incoming biases or the consequences of the researchers' interactions with participants, have been acknowledged and, wherever possible, sensitively addressed. 3. Direct quotations and excerpts from field notes are used to support the study findings, and enough of them have been included to inspire trust in the researchers' interpretations. 4. Multiple types of evidence have been used to inform conclusions. Findings suggest that the different types of evidence have been compared to check and confirm the identified patterns or themes. 5. It is easy to understand why and how the researchers did what they did, as well as see why they reached the specific conclusions outlined in the research report. Extensive information about the study's design, preparation, data collection, analysis, and results has been included.
Content analysis	Content analysis is a methodology for examining the content of something, such as instruction (e.g., how much instructional time is devoted to vocabulary instruction) or texts (e.g., what kinds of text, and in what proportions, are included in basal readers). Content analysis is more about the <i>what</i> in language, whereas discourse analysis is more about the <i>how</i> with language.	<ol style="list-style-type: none"> 1. The purpose of and theory and rationale for the content analysis is clear. 2. The sample and context from which the content is drawn is thoroughly described. 3. The variables for the content analysis are clearly defined and operationalized, with reference to the theoretical framework and rationale for the study. 4. There are at least two coders, who demonstrate high levels of coding agreement (inter-rater reliability). 5. The procedures for going from codes to results are reported clearly.
Correlational data	This kind of research examines relationships among variables. Researchers often conduct these studies when they are interested in causes and effects but are unable to control or alter the variables. For example, correlational research might examine the relationship between exposure to lead paint and reading difficulties.	<ol style="list-style-type: none"> 1. Concepts are carefully defined. 2. Statistical techniques are appropriately and widely used in the study. 3. Thoughtful attempts to identify and/or rule out alternative explanations have been made. 4. The findings from other related studies and existing theory are taken into account. 5. Any claims about cause and effect have been made with extreme care.
Discourse analysis	This methodology tries to gain insight into the structures and meanings that underlie conversations and written texts. Researchers examine previously or newly recorded texts and develop systems for uncovering patterns in the texts.	<ol style="list-style-type: none"> 1. Evidence that the data has been systematically collected and carefully analyzed has been provided. 2. Researchers have provided plausible and thorough explanations for why particular texts have been selected as focal texts and how the coding system was developed to analyze them. 3. Researchers' interpretations have taken into account the original contexts in which the texts were created, as well as existing research and theory. 4. High scores from the inter-rater reliability process have been reported. 5. Many direct quotations have been used to support conclusions.

(continued)

10 THINGS EVERY LITERACY EDUCATOR SHOULD KNOW ABOUT RESEARCH

Table 1 Brief Descriptions and Standards of Quality for Some Common Research Methodologies (continued)

Methodology	Description	Five quality indicators
Ethnography	Ethnography is a specific type of case study. Like case study research, ethnographic research explores phenomena by looking closely at specific examples. This kind of research typically involves extended, intense observations and emphasizes cultural contexts. Researchers often attempt to represent the perspectives of insiders.	<ol style="list-style-type: none"> 1. A large amount of time (often years) has been spent interacting with the participants. 2. Multiple types of evidence and evidence from different settings have been collected. 3. Researchers have carefully described the settings and circumstances of participants' lives, as well as their own incoming perspectives and the roles they enacted during the study. 4. Previous research and existing theory have been used to design the study and interpret the findings. 5. The study goes beyond providing new information about the focal topic by adding to and/or changing theory.
Experimental and quasi-experimental research	These designs investigate cause–effect relationships. Researchers typically identify a focus, such as the use of a particular instructional approach, and measure its outcomes. Researchers attempt to eliminate alternative explanations for outcomes by creating groups of participants who differ in only one way—for example, in receiving or not receiving a particular instructional approach. In experiments, researchers typically create groups by randomly assigning participants. In contrast, researchers use groups that already exist for quasi-experiments.	<ol style="list-style-type: none"> 1. Groups are highly similar, ideally differing only in terms of the variable of interest (e.g., receiving a treatment or not). 2. Researchers have tried to ensure that the research is valid and reliable (e.g., scoring assessments without knowing whether they are from the experimental or control group). 3. Appropriate and high-quality measures have been used to assess outcomes. 4. The statistical analyses are appropriate and clearly linked to the study design. Information that might affect readers' interpretation of the analyses, such as missing data or violations of the assumptions of particular statistical procedures, has been discussed. 5. Enough details about each group's participants and their study experiences have been included to compare participants and tasks to other studies and real-world situations.
Formative and design experiments	In this methodology, data are collected systematically for the purpose of informing design or practice to reach specified goals. Often, researchers and teachers work together to implement an instructional approach, investigate factors that might influence its outcomes, modify the approach to account for what they have discovered, and implement the revised instructional approach. This implement-investigate-and-revise process might continue for several rounds or until the original goal is achieved.	<ol style="list-style-type: none"> 1. The starting point (e.g., students' initial achievement) has been carefully explained. 2. Previous research and existing theory have informed the experiment. 3. Multiple rounds of data have been collected over time. 4. Revisions to the experimental design take into account the outcomes and circumstances of previous rounds of data collection. 5. The outcomes of each round are thoroughly discussed and include unplanned results.
Historical research	In historical research, researchers attempt to address a question about the past. They examine artifacts from or about the time period, such as diaries, photographs, court records, or legal documents. Researchers may also interview people associated with the event or topic. This kind of research often searches for patterns or themes that might inform current issues. For example, a researcher might examine past educational policies for the purpose of revising or creating present-day policy initiatives.	<ol style="list-style-type: none"> 1. The settings and circumstances of the time period are carefully described and are used to understand the study findings. 2. The research focus is precise and thoughtful. 3. Researchers describe what kind of and how much evidence is available to study the topic. Throughout the research report, they extensively use the available evidence and justify how and why particular sources were selected. 4. Researchers employ widely accepted and systematic methods to collect and analyze the data. 5. The limitations of the study are thoroughly discussed.
Neuroimaging	This kind of research tries to answer questions about neurological structure and/or function. Researchers examine images of the brain and brain activity. These studies are characterized by the use of specialized medical equipment and processes, such as electroencephalography or functional magnetic resonance imaging.	<ol style="list-style-type: none"> 1. Research questions can be addressed appropriately by looking at images of the brain and the brain's activity. 2. The tasks that participants are asked to complete seem suitable and appear to be high in quality. 3. The methods used to collect evidence are appropriate and straightforward. 4. Multiple approaches have been used to verify the findings. 5. Attempts have been made to compare and connect the findings with other research.

(continued)

10 THINGS EVERY LITERACY EDUCATOR SHOULD KNOW ABOUT RESEARCH

Table 1 Brief Descriptions and Standards of Quality for Some Common Research Methodologies (continued)

Methodology	Description	Five quality indicators
Quantitative meta-analysis	Researchers use this methodology to synthesize the results of previous research. Researchers systematically collect studies that have addressed the same or similar questions, then conduct statistical analyses to identify trends across the collected studies. Quantitative meta-analyses often focus on the relative magnitudes of outcomes, such as the average effect of a particular instructional intervention.	<ol style="list-style-type: none"> 1. The study focuses on an area of research that can support a quantitative meta-analytic approach. Enough research reports in the focal area include effect sizes or information that could be used to calculate effect sizes. 2. The search for previous research is systematic and thorough. Researchers have used multiple sources to obtain studies and have considered both published and unpublished works. 3. The criteria used in the quantitative meta-analysis seem thorough, appropriate, and comprehensive enough to lead to the identification of most or all relevant studies. 4. Appropriate statistical techniques are widely used in the study design. 5. High levels of inter-rater reliability or agreement have been reported.
Single-subject experimental design	In this design, individuals are (or an individual is) studied in such a way that they each comprise their own comparison group. For example, in an ABA withdrawal design, repeated baseline assessments are administered (A), then an intervention is introduced and the subject assessed repeatedly again (B), and finally, the intervention is withdrawn and the subject assessed additional times (A). Differences in A and B suggest a possible impact of the intervention.	<ol style="list-style-type: none"> 1. The focus of the measures/assessment is believed to be susceptible to intervention within a relatively short time frame. 2. Sensitive and reliable measures that can be administered repeatedly are used. 3. Participants and study procedures are described in great detail so that others can replicate them. 4. The design is such that any effect observed can be reasonably attributed to the intervention. 5. Any effects observed are evaluated for how much practical difference they make for the individual.
Survey research	Survey research usually elicits reports from participants about themselves. The purpose of this kind of research is usually to understand something about the larger group to which the participants belong. For example, researchers might survey 100 kindergarten teachers across a state to learn about the beliefs of the kindergarten teachers in that state. This kind of research may involve different kinds of interactions, such as face-to-face or telephone interviews and computerized or mailed surveys.	<ol style="list-style-type: none"> 1. The participants have been carefully chosen. Individuals are clearly representative of the larger group in which the researchers are interested. 2. The survey questions are straightforward and unbiased. 3. The quality of the survey has been checked, and evidence of its validity and reliability has been reported. 4. Reasonable attempts have been made to get participants to respond, as well as to understand why nonresponding participants did not respond. 5. Information that might influence readers' interpretations of the study, such as how the participants differ from the targeted population and what percentage of participants responded, has been included.
Verbal protocols	Verbal protocols, also referred to as think-aloud studies, typically gather information about people's thought processes. Researchers often ask participants to complete a specific task, such as reading a book, and report what they are thinking. Participants from second grade to adulthood have participated in verbal protocol research.	<ol style="list-style-type: none"> 1. The texts and tasks used in the study seem to be good choices for gathering information about the research questions. 2. The instructions and prompts are carefully worded to avoid influencing participants' responses. 3. The procedures for collecting data, including the prompts used, have been thoroughly described. 4. Plausible and logical reasons have been given for each research design decision. 5. Clear and logical connections can be seen among the research questions, research procedures, tasks, and materials.

Note. The information in this table was drawn in part from "Standards for Reporting on Empirical Social Science Research in AERA Publications: American Educational Research Association," by P.A. Moss, J.W. Pellegrino, B.L. Schneider, R.P. Duran, M.A. Eisenhart, F.D. Erickson, et al., 2006, *Educational Researcher*, 35(6), 33–40; "Qualitative Analysis on Stage: Making the Research Process More Public," by V.A. Anfara, Jr., K.M. Brown, and T.L. Mangione, 2002, *Educational Researcher*, 31(7), 28–38; *Literacy Research Methodologies*, by N.K. Duke and M.H. Mallette (Eds.), 2004, New York: Guilford; *Literacy Research Methodologies* (2nd ed.), by N.K. Duke and M.H. Mallette (Eds.), 2011, New York: Guilford; and *Educational Research: An Introduction* (8th ed.), by M.D. Gall, J.P. Gall, and W.R. Borg, 2007, Boston: Allyn & Bacon.

holders to continue to conduct or to ever have conducted research.

For some positions, a person may only need to have conducted a single research study: the dissertation study that originally earned the PhD or EdD. Of course, in many cases, professors are expected to and/or do in fact conduct a large number of other research studies, but that is by no means a given. In light of the range of roles that many professors play, they may at one moment be conducting a research study and at another moment be providing advice that as yet has little or no basis in research.

A flip side of these points is that just because someone is *not* a university-based researcher or does not have a PhD or EdD does not mean that he or she has not conducted research. Quality research can be and is conducted outside of a university context: in educational research foundations, in schools or other educational settings, and in community centers. Quality research can be and is conducted by individuals who do not have a research degree. Similarly, it is possible for people who are not researchers and not in university positions to have a strong command of research even if they do not conduct it themselves.

4. The Difference Between Research-Based and Research-Tested

When invoking the term *research* in relation to instructional practices, approaches, or products, we think it is valuable to make a distinction between *research-based* and *research-tested*. (As will become evident, this is both similar to and different from Shanahan's [2002] distinction between *research-based* and *research-proven*.) We use *research-tested* to mean that one or more research studies tested the impact of that particular

practice, approach, or product. In contrast, we use *research-based* to mean that the particular practice, approach, or product has not been tested in a research study but has been designed to be consistent with research findings.

Let's say that someone develops an approach to teaching reading comprehension that combines elements of three previous, research-tested approaches. We would say that the new approach that has been developed is research-based, but we would not say it is research-tested because no one has tested the impact of that particular combination of approaches in research. The combination could be more effective, as effective, or less effective than the individual approaches on which it was based.

The terms *research-tested*, *research-based*, *research-proven*, and others are being used by many to elevate the status of their product or approach. However, use of these terms alone means very little. We have to ask questions that allow us to get underneath any individual's, organization's, or company's use of these terms. In the remainder of this section, we identify some of the questions that need to be asked when a product or approach is said to be research-tested or research-based.

Research-Tested

To say that a practice, approach, or product is research-tested, or research-proven, sounds like a powerful endorsement or attribution, but its strength really

"We think it is valuable to make a distinction between research-based and research-tested."

depends on how it was tested and what the tests found. Questions we need to ask include the following:

■ *What exactly did the research test?*

For the research-tested claim to be valid, it needs to be applied to a practice, approach, or product that is highly similar to or the same as what was originally tested.

■ *What exactly did the research find?*

Just because something was research-tested does not mean that the test/research found it to be effective.

■ *Did the research test the practice, approach, or product against something else?*

Simply trying out something provides limited information. If the practice, approach, or product is not tested against something, then it is difficult to infer whether any impact that we observe would have been seen even without that practice, approach, or product. For example, students may appear to have better reading comprehension after receiving a particular kind of instruction, but it may well be the case that their reading comprehension is simply developing as it normally would have without the intervention or that something else entirely is causing the observed improvement. For these reasons, having a comparison group that differs only in that they did not use the particular practice, approach, or product is important to addressing questions about impact.

■ *To what exactly was the practice, approach, or product compared?*

When there is a comparison, you need to ask what it is. It is one thing for a practice, approach, or product to be effective when compared with recess or completing worksheets, and quite another to show that it is effective in comparison to good, quality instruction designed to address that same area of learning.

“To say that a practice, approach, or product is research-tested, or research-proven, sounds like a powerful endorsement...but its strength really depends on how it was tested and what the tests found.”

■ *With what sample(s) was the research conducted?* We have to be cautious about applying the results of a study with one population to another population if we think the populations are different in consequential ways. For example, if an approach has only been tested with native English speakers, then its use with English learners should be approached with caution. Similarly, an approach tested and shown to be efficacious in only small, homogeneous classes may not be as efficacious in large, heterogeneous classes. Ideally, there should be a close match between the sample or samples with which the practice, approach, or product has been tested and the group with whom you plan to use it.

■ *What outcome measures were used?* The outcome measures used in a study should be consistent with the outcomes you are concerned about for your particular purpose or context. For example, suppose your goal is to improve children’s fluency, or accuracy, rate, and prosody, in reading connected text. Suppose you find a study in which an intervention improved students’ accuracy and rate at reading lists of nonsense words but did not improve the accuracy and rate with which students read connected text. In such a case, you might look for other studies or interventions with an outcome measure aligned more closely to your goals. When there is a mismatch between the outcome measures of a study and what you are trying to impact, the fact that the practice,

approach, or product proved to be effective may be beside the point.

■ *What impact did the research find?* Another question to ask is about the degree of the impact of a particular practice, approach, or product. If a practice requires the input of considerable resources to implement and results in children learning 1.6 more alphabet letters on average, then it may not be worth the investment. In contrast, if a practice required relatively few resources to implement and improved children’s spelling development by several stages in a short period of time, then it may well be worth the investment.

■ *How many studies were conducted?* A practice, approach, or product that has been tested and shown to be effective in dozens of research studies with a wide range of contexts should be viewed as having stronger research support than a practice, approach, or product that has been tested in only a single study or in a very narrow range of contexts. Similarly, all things being equal, a practice, approach, or product that has been shown to be effective in some studies and not others should be seen as having less research support than one that has been shown to be effective in every study in which it was examined.

■ *What was the quality of the studies?* You also need to ask whether the research meets standards of quality for that type of research. We talk more about standards of quality for different types of research later in this article (for

research-tested claims, see especially “Experimental and quasi-experimental research” in Table 1).

Research-Based

Just as we must ask questions to get underneath claims that something is research-tested or research-proven, so too must we ask questions to get underneath claims that something is research-based. Questions that we should ask include the following:

- How many studies were conducted, and what were the findings?
- How similar is what was researched and found to what was being given the research-based label?
- What research design(s) was(were) used, including the type of research, the sample, outcome measures, and so forth?
- Of what quality were the studies?

Notably, we should not assume that something that is research-tested is inherently more supported by research than something that is research-based. For example, a practice that has been tested and found to be effective in a single study (i.e., research-tested) is, in our view, generally less compelling than a practice that is similar to, but not exactly the same as, a number of practices that have been tested and shown to be effective in a larger number and wide range of studies (i.e., research-based).

5. Many Kinds of Research Have Valuable Contributions to Make to Our Understanding of Literacy Learning, Development, and Education

You may have read references to the experiment or randomized, controlled

trial as the “gold standard” in educational research. Some seem to think that experiments—studies in which individuals or groups are randomly assigned to different experiences and then the results are compared—are the best and most valuable kind of educational research. We believe that this way of thinking is mistaken and misleading. Instead, many kinds of research have valuable contributions to make.

For example, research examining the validity and reliability of an assessment is valuable. Survey research telling us how motivated, or unmotivated, U.S. students are to read is valuable. Research following a single struggling reader in and out of classrooms over a period of years is valuable. Research analyzing the characteristics of young children’s spelling is also valuable. The educational enterprise is far too complex for one type of research to answer all of our questions or meet all of our needs (e.g., Shavelson & Towne, 2002).

6. Different Kinds of Research Are Good for Different Questions

Although many different kinds of research are valuable, they are not valuable for the same reasons. For example, an experiment is a valuable tool in learning whether one approach to instruction is more or less effective than another, but an experiment is not particularly well suited for understanding what is going on in the minds of good readers as they

read; verbal protocols and neuroimaging would work well for that.

Similarly, although instrument development is of great value when trying to understand how to assess or measure something, it does not, alone, help us gain in-depth understanding of the social and cultural context of a particular group of learners. The type of research design that is most valuable depends on the research question that one is trying to address (Shavelson & Towne, 2002, 2004).

7. High-Quality Research Has a Logic of Inquiry

According to Wilkinson and Bloome (2008),

What matters in the evaluation of the worth of a piece of research, of any paradigm or intellectual tradition, is the manner in which researchers locate their inquiry against a background of extant knowledge and assumptions, the goodness of fit between research questions and methodologies, the quality of the data collection and analysis, and the integrity of the overall warrant for the claims (Howe & Eisenhart, 1990). (p. 7)

In other words, a good research study has a strong match between the research question or purpose, the research design, and the conclusions drawn and claims made from the research. For instance, consider the question, What do highly effective literacy teachers do during small-group reading instruction? To address the question, the researcher must develop a defensible means of identifying effective

literacy teachers, a systematic procedure for collecting data about their practices during small-group reading instruction (e.g., observation, audiotaping), and a systematic procedure for analyzing the data (e.g., coding it for specific kinds of practices, generating descriptive statistics to identify common practices).

When this is complete, the researcher has to take care not to make claims that cannot be substantiated by the data that has been collected. For example, the claim “Telling readers to identify the selection’s text structure prior to reading is a highly effective instructional practice” is not warranted given the design. Even if all of the observed teachers engaged in this practice, we cannot be sure that this is what makes their instruction effective. In contrast, the claim “Highly effective teachers regularly tell readers to identify the selection’s text structure” might be well supported by the design.

8. Conclusions Drawn From Research Are Only as Sound as the Research Itself

It is extremely important to look at the quality of the research design and implementation when evaluating a particular claim or conclusion. Conclusions based on a seriously flawed study may be seriously flawed. A logic of inquiry, discussed previously, must be evident in the research. Also, the nature of the inquiry will call for different markers of quality. In Table 1, we present some standards of quality for several kinds of research. For further information about evaluating the quality of any of these kinds of research, we suggest consulting methods texts, such as *Fundamentals of Educational Research* (Anderson, 1998), *Educational Research: An Introduction* (Gall, Gall, & Borg, 2007), or *Literacy*

“The educational enterprise is far too complex for one type of research to answer all of our questions or meet all of our needs.”

Research Methodologies (Duke & Mallette, 2004, 2011).

9. Where and How Research Is Published or Presented Requires Particular Attention

Consider a particular news item and the range of different ways it is covered, for example, by the *New York Post*, *The New York Times*, *Newsweek*, *The Economist*, Fox News, or the MacNeil/Lehrer News Hour. These sources will cover the same story in substantially different ways. Similarly, literacy research in different outlets, and by different writers, may be reported very differently.

For example, a study reported in *The Reading Teacher* may provide relatively little detail about a study's methodology and findings but may provide considerable detail about the targeted instructional practices; in contrast, a study reported in *Reading Research Quarterly* may provide much more detail about the study's methodology and findings but may provide a much briefer explanation of how to implement the given instructional practices (cf. Duke, Purcell-Gates, Hall, & Tower, 2006; Purcell-Gates, Duke, & Martineau, 2007). A study reported to an audience primarily comprised of literacy educators may give short shrift to institutional, fiscal, and equity issues—considerations that would likely be prominent in a report of the same study intended for an audience primarily comprised of education policymakers (cf. Buly & Valencia, 2002; Valencia & Buly, 2004).

As you read or listen to research or references to research, it is important to consider the match or mismatch between your purpose(s) and the purpose(s) for which a given piece was written or presented in that particular outlet and by

“Consider the match or mismatch between your purpose(s) and the purpose(s) for which a given piece was written or presented.”

that particular person. If your intent, for instance, is to understand and evaluate the quality of the methodology used in the study, then an article in *Instructor Magazine* is not the best place to go; however, if your intent is to understand what a particular research-tested practice looks like in a school setting, then this magazine might be the correct source. In the end, it would be best to read everything you can, as presented in a range of outlets, about any specific study or area of research that is of great importance to you.

You may have heard or read about the notion of peer-reviewed research. This refers to whether the research was reviewed by a set of peers prior to its publication. Publications like *The Reading Teacher* and *Reading Research Quarterly* generally, although not always, employ a blind peer-review process. This means that when deciding whether to publish a particular submission, the editors send it out to a set of scholars (researchers in the case of *Reading Research Quarterly*, and a combination of researchers, practitioners, and other scholars in the case of *The Reading Teacher*) for their evaluation about whether it should be published and, if it is worthy of publication, to find out what improvements could be made to the piece prior to publication.

The adjective *blind* means that submitted articles are sent to reviewers without

information that would clearly identify the submission's author(s); this increases the likelihood that reviewers are not swayed in their judgment by their high or low opinions of the submitted article's author(s). The intent of peer review is to inform decisions about what merits publication and improve the quality of what is ultimately published. Of course, this tool is not perfect. For various reasons, work of dubious quality is sometimes published, and sometimes high-quality research is not. A piece's peer-reviewed status is just one metric that you can use to increase the likelihood that the research you are reading is of high quality, but you should keep in mind that it is just one, necessarily limited, metric.

Another issue to consider with respect to how the research is presented is the quality of the writing itself. All other things being equal, and given that our time for reading research is always limited, it is worth considering the quality of a piece's writing in deciding whether to read it. That said, all other things are often not equal. A study may follow a sound logic of inquiry as described previously, meet standards of quality as described in Table 1, and have many other markers of high quality but be poorly written. For example, important information may be missing, or the piece's prose may be dense and difficult to parse. Conversely, very strong writing may make a study sound better than it actually is. For example, the piece's prose may be so articulate and persuasive that it is easy to miss that its design fails to meet many standards of quality.

These are important considerations to bear in mind as you read research. It may be worth plowing through an article that is difficult and frustrating to read, and we have to be careful not to be seduced by a study that, although

presented eloquently, is problematic in important ways.

10. Educational Research Proceeds Through the Slow Accumulation of Knowledge

When designing a study, there are typically trade-offs. For example, involving more classrooms may increase the generalizability of the study—that is, the extent to which you can apply what is learned in the study to other classrooms—but may decrease our capacity to study in detail how individual teachers are implementing the intervention. Even if there were a perfect study (and of course, researchers dream of conducting a perfect, flawless study!), we must be cautious about basing policy or practice on any *single* study.

First, it is possible for any study, no matter how well designed, to get an anomalous result. Second, the generalizability of any one study is limited in terms of the populations and contexts in which similar results can be expected to be found. For these reasons, it makes the most sense to think about research as proceeding as a slow accumulation of knowledge over time and to read across many different studies on a particular question or topic. Although breakthroughs or headline-making studies periodically appear, it is usually a mass of related studies over a period of years that lead to a well-accepted or durable conclusion.

An important tool that can help us understand whether practices or insights have been widely validated is the research review or synthesis. These syntheses pull together individual studies to draw more robust conclusions. These syntheses may also be a great asset for educators who are too busy to read a lot of studies.

There have been a number of influential research reviews in literacy education in recent years, most notably, perhaps, the National Reading Panel report (National Institute of Child Health and Human Development, 2000) but also several others, such as *Preventing Reading Difficulties in Young Children* (Snow, Burns, & Griffin, 1998), *Writing Next: Effective Strategies to Improve Writing of Adolescents in Middle and High Schools* (Graham & Perin, 2007), the National Literacy Panel report on language-minority children and youths (August & Shanahan, 2008), and the National Early Literacy Panel report (National Early Literacy Panel, 2008).

There are also volumes that compile many research reviews or syntheses, as in the *Handbook of Reading Research* (e.g., Kamil, Mosenthal, Pearson, & Barr, 2000; Kamil, Pearson, Moje, & Afflerbach, 2011) and *What Research Has to Say About Reading Instruction* (e.g., Farstrup & Samuels, 2002; Samuels & Farstrup, 2011). Finally, there are also outlets for peer-reviewed reviews of research, such as the journal *Review of Educational Research*.

However, these important, and potentially time-saving, tools also must be treated with caution. They are necessarily reductive. They may omit important details about the methods used to conduct the studies or about findings of the studies, particularly those findings that are not directly germane to the focus of the review. These publications may also gloss over important differences between studies, the ways particular items have been operationalized, and so forth. Ideally, you would go back to all of the original papers on which the review is based, as well as search for and read any relevant papers that may have been omitted. However, this is often not possible. If

it is not, then there are other questions that you could ask:

- Are the authors of the review well respected in the field?
- Do they represent, or at least show openness to, a variety of theoretical perspectives related to the research reviewed?
- Does it appear that the authors are indeed reviewing all of the available research rather than simply selecting research that is consistent with their conclusion or perspective?

Was their approach to locating and reviewing the research systematic? Did their approach seem thorough?

If provided, was the list of terms used in the literature search relevant and exhaustive?

- Did another person also code the studies? To what degree did he or she agree with the authors' coding efforts?
- If you are familiar, or can make yourself familiar, with an individual study or a few studies reviewed, did the authors code, characterize, and draw conclusions that you think are appropriate from this study?
- Are the authors explicit in their explanations of findings and results?

Additional standards apply in the case of a special kind of research review called a quantitative meta-analysis (see Table 1 for more information). Quantitative meta-analyses can be very valuable, but it is important to note that these analyses can only examine studies that are correlational, quasi-experimental, or experimental in design, and these analyses are limited to only those areas in which a critical mass of

sufficiently similar studies have been conducted. For this reason, it is often not a quantitative meta-analysis—or not solely a quantitative meta-analysis—that will best address your needs.

Conclusion

By this point, it is likely clear that being an informed and critical reader of research is a formidable task. Thus, we end this article the way we began it, with the assertion that research is worth it—that research should be seen as an essential guide to policy and practice. Recall our argument that our experiences alone may misguide us, that sometimes we do not know what we do not know, that research allows us to take a longer term view than our personal experiences may allow, that research allows us to pool our numbers and experiences, and that research allows us into places and situations that we may not be able to observe otherwise.

Recall that the goal of literacy researchers is much the same as the goal of literacy educators: to improve literacy teaching and learning. Literacy educators play a vital role in helping literacy researchers meet this goal. Careful reading, evaluation, and interpretation by thoughtful and informed educators offers our best chance at realizing the full value of what this priceless tool—research—has to offer.

REFERENCES

- Anderson, G. (with Arsenault, N.). (1998). *Fundamentals of educational research* (2nd ed.). Bristol, PA: Falmer.
- Anfara, V.A., Jr., Brown, K.M., & Mangione, T.L. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher*, 31(7), 28–38. doi:10.3102/0013189X031007028
- August, D., & Shanahan, T. (Eds.). (2008). *Developing reading and writing in second-language learners: Lessons from the report of the National Literacy Panel on language-minority children and youth*. New York: Routledge; Washington, DC: Center for Applied Linguistics; Newark, DE: International Reading Association.
- Bos, C.S., & Anders, P.L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly*, 13(1), 31–42.
- Buly, M.R., & Valencia, S.W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis*, 24(3), 219–239. doi:10.3102/01623737024003219
- Duke, N.K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly*, 35(2), 202–224. doi:10.1598/RRQ.35.2.1
- Duke, N.K., Bennett-Armistead, V.S., & Roberts, E.M. (2003). Bridging the gap between learning to read and reading to learn. In D.M. Barone & L.M. Morrow (Eds.), *Literacy and young children: Research-based practices* (pp. 226–242). New York: Guilford.
- Duke, N.K., & Kays, J. (1998). “Can I say ‘once upon a time’?”. Kindergarten children developing knowledge of information book language. *Early Childhood Research Quarterly*, 13(2), 295–318. doi:10.1016/S0885-2006(99)80041-6
- Duke, N.K., & Mallette, M.H. (Eds.). (2004). *Literacy research methodologies*. New York: Guilford.
- Duke, N.K., & Mallette, M.H. (Eds.). (2011). *Literacy research methodologies* (2nd ed.). New York: Guilford.
- Duke, N.K., Purcell-Gates, V., Hall, L.A., & Tower, C. (2006). Authentic literacy activities for developing comprehension and writing. *The Reading Teacher*, 60(4), 344–355. doi:10.1598/RT.60.4.4
- Farstrup, A.E., & Samuels, S.J. (Eds.). (2002). *What research has to say about reading instruction* (3rd ed.). Newark, DE: International Reading Association.
- Gall, M.D., Gall, J.P., & Borg, W.R. (2007). *Educational research: An introduction* (8th ed.). Boston: Allyn & Bacon.
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools—a report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Guthrie, J.T., Wigfield, A., & Perencevich, K.C. (Eds.). (2004). *Motivating reading comprehension: Concept-oriented reading instruction*. Mahwah, NJ: Erlbaum.
- Kamil, M.L., Mosenthal, P.B., Pearson, P.D., & Barr, R. (Eds.). (2000). *Handbook of reading research* (Vol. 3). Mahwah, NJ: Erlbaum.
- Kamil, M.L., Pearson, P.D., Moje, E.B., & Afflerbach, P. (Eds.). (2011). *Handbook of reading research* (Vol. 4). New York: Routledge.
- Moss, B., & Newton, E. (2002). An examination of the informational text genre in basal readers. *Reading Psychology*, 23(1), 1–13. doi:10.1080/027027102317345376
- Moss, P.A., Pellegrino, J.W., Schneider, B.L., Duran, R.P., Eisenhart, M.A., Erickson, F.D., et al. (2006). Standards for reporting on empirical social science research in AERA publications: American Educational Research Association. *Educational Researcher*, 35(6), 33–40.
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Pappas, C.C. (1993). Is narrative “primary”? Some insights from kindergartners’ pretend readings of stories and information books. *Journal of Reading Behavior*, 25(1), 97–129.
- Perry, K.H. (2009). Genres, contexts, and literacy practices: Literacy brokering among Sudanese refugee families. *Reading Research Quarterly*, 44(3), 256–276. doi:10.1598/RRQ.44.3.2
- Purcell-Gates, V., Duke, N.K., & Martineau, J.A. (2007). Learning to read and write genre-specific text: Roles of authentic experience and explicit teaching. *Reading Research Quarterly*, 42(1), 8–45. doi:10.1598/RRQ.42.1.1
- Samuels, S.J., & Farstrup, A.E. (Eds.). (2011). *What research has to say about reading instruction* (4th ed.). Newark, DE: International Reading Association.
- Shanahan, T. (2002). What reading research says: The promises and limitations of applying research to reading education. In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 8–24). Newark, DE: International Reading Association.
- Shavelson, R.J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- Shavelson, R.J., & Towne, L. (2004). What drives scientific research in education? Questions, not methods, should drive the enterprise. *Observer*, 17(4). Retrieved August 7, 2009, from www.psychologicalscience.org/observer/getArticle.cfm?id=1557
- Snow, C.E., Burns, M.S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Snowling, M.J. (2000). *Dyslexia* (2nd ed.). Malden, MA: Blackwell.
- Valencia, S.W., & Buly, M.R. (2004). Behind test scores: What struggling readers really need. *The Reading Teacher*, 57(6), 520–531.
- Wilkinson, I.A.G., & Bloome, D. (2008). Research as principled, pluralistic argument. *Reading Research Quarterly*, 43(1), 6–8.