Research and Innovation: Let the Buyer Beware

Dr. J.E. Stone, East Tennessee State University
Andrea Clements, East Tennessee State University

Originally published in The Superintendent of the Future. (1998)
Robert R. Spillane & Paul Regnier (Eds.). Gaithersburg, MD: Aspen Publishers.

Can be ordered at at www.aspenpublishers.com

Provided courtesy of:

Education Consumers Foundation

1655 North Fort Myer Drive, Suite 700; Arlington, VA 22209
Phone (703) 248-2611; Fax (703) 525-8841
ecf@education-consumers.org

www.education-consumers.org

Research and Innovation: Let the Buyer Beware

Dr. J.E. Stone and Andrea Clements, East Tennessee State University

Schools are inundated with research that promises to improve achievement. Yet when programs are implemented results always seem to fall short. How can it be in school after school, year after year? The answer depends on whom you ask.

Educational researchers allege that the problem is a lack of money for research. They say that research and development has little impact simply because there isn't enough of it and they cite studies showing that funding is meager relative to the magnitude of the education enterprise. Outside observers disagree. Some 35,000 professors of education at American colleges and universities devote an average 14% of their time to research--broadly defined. Their students conduct research too. Annually, more than 7,300 doctoral students in education write dissertations. Myron Lieberman (1993) estimates the dollar value of the manpower dedicated to educational research by professors and doctoral students alone to be in excess of $700 million annually. Still other education research is authored by state departments of education, by nonprofit "think tanks," by federal agencies, and by the regional educational research laboratories. Significantly, only a small percentage of published research is undertaken by schools or school systems.

The results of this scholarly activity are readily available to schools through a variety of sources. Thousands of books, professional and academic journals, newsletters, technical bulletins, and other published sources make research available to teachers and administrators. Many recent publications are available on the Internet. A vast amount of material is indexed in the federally sponsored Education Resources Information Center (ERIC). ERIC includes a Current Index to Journals in Education and a microfiche library of mostly unpublished research called Research in Education. Research in Education is available in education libraries throughout the United States. The amount of research available through these several sources is staggering, and most of it is directly or indirectly related to the problem of improving school achievement.

The idea of improving teaching through the application of science has been around since the earliest days of organized teacher training. John Dewey, for example, believed that the scientific study of child development would improve classroom instruction by suggesting ways in which teaching might be fitted to

the learner (Dewey, 1916/1963). However, it was not until the 1960s that governmentally funded research began expanding to present-day levels. The Johnson administration's "war on poverty" infused federal dollars into university research institutes and education laboratories on an unprecedented scale. Head Start (U.S. Department of Health and Human Services, 1985) and Follow Through (Proper & St. Pierre, 1980) are prime examples. Both were designed to improve the school success of disadvantaged children and they are among the largest educational research projects ever mounted. The Follow Through project alone cost nearly $1 billion.

Has the money and manpower spent on research been justified by improvements in schooling? If the findings reported in Education Week's "Quality Counts" (Wolk, 1997) are any indication, the answer would have to be no. Despite the pressures for improvement created by reports such as the National Commission on Excellence in Education's A Nation at Risk (1983), measured achievement has stayed essentially flat. The National Assessment of Educational Progress scores in math and science have risen only a few points on a 500-point scale since 1973 (U.S. Department of Education, 1996). Of course there are isolated examples of significant improvement, but the broad picture is that the schools are (in the words of "Quality Counts") "treading water."

WHY SO LITTLE IMPACT?

If there is a significant amount of research--although arguably not enough--and the findings are widely available, why is there not at least a trend toward improved achievement? Again, researchers have an answer: Good research is available but schools fail to implement it. In other words, schools talk as though they adopt research-based innovations but at the classroom level they keep doing the same old thing (Cuban, 1993). There is more than a little truth to this claim. The innovative programs publicized by school administrators are not always translated into classroom practice. Teachers have a great deal of independence in the classroom and they are taught to fit their teaching style to students' needs. Remaining with accustomed approaches is, indeed, the tendency if only for reasons of comfort and familiarity.

Another explanation offered by researchers is that schools don't know good research when they see it. They are easily drawn to familiar practices supported by weak evidence. Unfamiliar practices supported by very credible evidence are often ignored. As discussed below, there is merit to this view. From the standpoint of

science, experimental studies are far more convincing than descriptive and correlational ones, yet school personnel often ignore the stronger and adopt innovations suggested by the weaker. For example, during the 1960s and 1970s correlational studies suggesting self-esteem enhancement as a means to improved achievement led to sweeping changes in teacher training and schooling. Experimental findings to the contrary were ignored (Scheirer & Kraut, 1979). They showed that self-esteem and achievement are correlated mainly because achievement enhances self-esteem, not because self-esteem enhances achievement.

One other explanation popular with researchers is the institutional inertia warps and retards progress. Plainly this view also has merit. All organizations encourage some possibilities and restrict others. All are comfortable with certain ways of conducting themselves and uncomfortable with others. Teacher unions, for example, may resist changes that make teachers' jobs more laborious. Administrative customs may resist change that make jobs look too easy. Of course, community expectations, regulatory policy, and public oversight can all exert resistance to change.

In marked contrast to the views of researchers, schoolhouse "insiders" (i.e., teachers and administrators) say that research has little impact because much of it does not work in the real world. As they see it, schools are doing everything they can to implement the latest findings, but social and economic realities impose limits. Implementing research is like rebuilding a ship in the midst of a voyage. Staying afloat has to be the first consideration. Rebuilding during a storm is even more problematic. Schools can and do make the changes suggested by research, but circumstances can trump even the best-laid plans. Even with successful implementations, effects are obscured or nullified by factors such as limited resources, two-earner families, increased crime, teen pregnancy, drug abuse, gangs, television, and a host of other hindrances and adversities (Olson, 1997).

Despite the often limited benefit of research-based innovations, schools continue to adopt them--if only to keep up with the latest trends. Which research and which innovations, however, often depends less on the quality of the findings than on the channel through which the research comes to the school's attention. School personnel are frequently exposed to "the latest" research at workshops, professional meetings, and in-service training. Typically, the teachers, administrators, and board members who attend these meetings have a limited understanding of research and/or of the findings pertaining to the innovation in question. More often than not, presenters and programs for such meetings are selected not because their ideas are well grounded but because they have a stimulating presentation. In addition, audience interest is often spurred by a regulatory mandate or incentive funding, not a burning desire for improved student achievement.

Other pragmatic considerations play a role as well. For example, attractiveness to students, teachers, parents, and other school system stakeholders can weigh heavily in research selections. So can public relations. For example, the desire of school leaders and board members to demonstrate "progressive leadership" often plays a contributory role. In short, the selection of research-based programs and innovations brought back from workshops and meetings may be substantially influenced by considerations other than evidence of effectiveness.

The Restrictions Imposed by Doctrine

Another factor that influences decisions about research is educational philosophy. The practice of injecting popular psychological theory into schooling--often without regard to effectiveness or applicability--has been a chronic problem in American education (Davis, 1943; Hilgard, 1939). Currently, a poorly recognized but longstanding educational doctrine called "developmentalism" (Hirsch, 1996; Stone, 1996) permeates the public schooling community. Developmentalism frames teaching and learning issues in a way that favors certain types of research and disregards others.

Developmentalism is a derivation of eighteenth-century romantic naturalism. The French philosopher Jean Jacques Rousseau (1712-1778) is the most influential of its early proponents. The works of John Dewey (1859-1952) and Jean Piaget (1896-1980), however, are more directly responsible for its present-day acceptance. Developmentalism is a view of age-related social, emotional, and cognitive change that presumes a child's native tendencies to be a fragile expression of the individual's natural and therefore optimal developmental trajectory (Stone, 1996). It conceives of education as a set of experiences that serves to facilitate and preserve that trajectory by fitting the educational experience to the individual.

Developmentalism contrasts sharply with the classic tradition in education and with the American tradition founded by the Puritans. Both sought to civilize and better the individual, not merely accommodate his or her emerging tendencies. Both classic tradition and the common school aimed to discipline natural impulses in service of a higher good. The significance of this philosophic issue as an impediment to effective schooling would be difficult to overstate. Most public schools seek achievement to the extent permitted by students' natural inclinations. They are "learner centered." Most parents and policy makers want schooling that impels achievement beyond that to which most students are inclined by their youthful proclivities (Steinburg, 1996). They are "learning centered."

The dominance of learner-centered pedagogy is in no small part an accident of history. Progressivism--a social and philosophical offshoot of romantic naturalism--predominated in American intellectual circles in the late nineteenth century and early twentieth century. These were the years during which universal public education came to be public policy as well as the formative years of many teacher-training institutions. Accepted teaching practices of that day were often harsh and punitive; thus progressive methods were a welcome alternative. The premier teacher-training institution of the early twentieth century was Teacher's College, Columbia University (Cremin, 1964). Its graduates led the development of other such programs around the country. Even today, the educational methodologies that prevail in the public education community are those that agree with the philosophic leanings of the Teacher's College faculty of the early 1900s (Hirsch, 1996).

Developmentally informed pedagogy has come to dominate public schooling but without clear public recognition of its nature and its role. Over the past 75 years it has emerged and reemerged under a variety of names. In the 1920s it was called "progressive" and "child centered." Today it is termed "reflective" and "learner centered" (Darling-Hammond, Griffin, & Wise, 1992). However termed, it has consistently maintained that teachers should seek to instruct only through activities that students find engaging and enjoyable. Thus, instead of employing the most enjoyable of teaching methods that are known to result in learning, teachers have been trained first to seek activities that are enjoyable and engaging and to use them in ways that will produce learning. Thus good teaching has come to be thought of as teaching that is well received and that incidentally produces some degree of learning.

Uncertainty about learning outcomes was not considered a pedagogic weakness by progressive education's founders. Neither John Dewey nor progressive education's great popularizer, William Heard Kilpatrick, considered conventionally prescribed educational objectives to be the proper aim of schooling. Instead, both argued that schooling should seek the emergence of an individually defined and broadly conceived intellectual development. Dewey, in particular, wrote at length about the harm done by teacher insistence on externally defined aims (Dewey, 1916/1963). Viewed from the progressive/learner-centered perspective, research that seeks to demonstrate a teaching methodology's ability to produce a preconceived learning outcome is inherently faulty and inconsistent with the proper aims of schooling.

Despite public repudiation in the 1950s, Dewey's view remains the foundation of today's cutting-edge innovations. It has spawned a remarkable array of educational terms and concepts, and they have been widely propagated by agencies and organizations such as the U.S. Office of Education, the state departments of

education, teacher-training programs, accrediting agencies, professional and academic societies, and the like.

The education community seeks to improve schooling through the use of research, but learner-centered strictures guide the adoption process. The impression created by the vast assortment of current educational terms and concepts is one of abundant variety. In truth, however, most conform to the same progressive vision of education. As noted by E. D. Hirsch (1996), "within the educational community, there is currently no thinkable alternative" (italics in the original, p. 69). Recent permutations and derivatives include the following:

- lifelong learning
- developmentally appropriate instruction
- brain-based learning
- situated learning
- cooperative learning
- multiple intelligences
- multiaged instruction
- discovery learning
- portfolio assessment
- constructivism
- hands-on learning
- project method
- thematic learning
- integrated curriculum
- higher-order learning
- authentic assessment
- whole-language reading

How Learner-Centered Thinking Restricts Choices: The Case of the Follow Through Project

Learner-centered doctrine discourages the use of results-oriented research (Stone, 1996). Studies concerned with improving achievement typically test an intervention or treatment (i.e., an action taken by the researcher that is intended to produce change in the student). The success of the intervention is judged in reference to some predetermined expectation. In contrast to the goal of inducing results, the goal of developmentally informed research is to accommodate schooling to the individual and to do so in a way that achieves the ends to which the individual is inclined by nature, not those prescribed by the curriculum.

One of the clearest instances of results-oriented research rejected on learner-centered grounds comes from

the Follow Through project (Proper & St. Pierre, 1980). Follow Through was a huge federally funded research project of the late 1960s and early 1970s. It was launched in 1967 by the Ninetieth Congress in response to President Johnson's request to "follow through" on project Head Start. Improved achievement in the basic skills of disadvantaged students was its prime objective. It remains the largest educational experiment ever.

Nine educational models were compared in 51 school districts over a six-year period. Of the nine, all but two were learner centered; and contrary to the prevailing educational wisdom, the two exceptions significantly outperformed the field. Of greater significance, five of the seven learner-centered models produced worse results than the traditional school programs (i.e., the nontreated control groups) to which each Follow Through approach was compared. What makes the contrast especially striking is that the outcome measures included not only basic skills but "higher-order" cognitive skills and a measure of self esteem--the very sort of outcomes that learner-centered methods are intended to enhance.

The most successful of the nine models was Direct Instruction (Engelmann, Becker, Carnine, & Gersten, 1988)a structured and so-called teacher-centered approach. Despite its overwhelming success, Direct Instruction was disparaged and largely ignored by the education community (Watkins, 1988). A lengthy critique of Follow Through was published in Harvard Educational Review (House, Glass, McLean, & Walker, 1978), and the U.S. Department of Education's National Diffusion Network--a bureaucratic agency responsible for disseminating only the "best" research--concluded that all nine programs were valid and all were recommended for further funding. In fact, added funding was given to the failed models on the grounds that they needed strengthening.

The Follow Through Direct Instruction findings are by no means the only research that has been ignored because it disagreed with the learner-centered view. Herbert Walberg (1990, 1992) summarized some 8,000 reports of demonstrably effective teaching methods. Like Direct Instruction, most were structured, teacher-directed, and designed to produce measurable gains in achievement. Most could be described as learning-centered instead of learner-centered. Many employed drill, recitation, and incentives for student effort. A review of research literature by Ellson (1986) found 75 studies of teaching methods that produced achievement gains at least twice as great as those of comparison groups. Many of them were popular at one time but none are learner-centered and none are in widespread use today.

The reception accorded Direct Instruction and other learning-centered research is important because it highlights a critical difference between the public's educational objectives and those of the learner-centered

schooling establishment. Public Agenda (Johnson & Immerwahr, 1994) and other public polling organizations have found that the public wants schools that produce conventionally measured academic achievement. The public is not opposed to the goals of learner-centered schooling, but it considers them secondary to conventional academic achievement. To the public, outcomes such as improved self-esteem are attractive, but schools that fail with respect to academic achievement are nonsense no matter what else they may produce. The same priorities are embodied in state-level school accountability policies. They focus primarily on academic gains operationally defined by achievement tests. By contrast, learner-centered research gives equal priority to "intellectual growth," enhanced self-esteem, and gains in knowledge and skills. If one or more of the three are produced, the research is taken to be informative and potentially valuable for school implementation.

Why Researchers Remain Learner-Centered

Despite the ever-growing demand for improved achievement, neither researchers nor schools are able to break away from learner-centered thinking, and for several reasons. Both researchers and most school personnel are indoctrinated in learner-centered thinking, and powerful incentives encourage them to remain loyal to that point of view.

For researchers, funding is a prime incentive. Fund allocations are almost inevitably influenced by other educators, and most of them subscribe to learner-centered orthodoxy. Funding affords a researcher time to work, and to have a reasonable chance at funding, one's proposal must appeal to the views of other educators.

For most researchers, funding is tied to institutional support. Most researchers are college faculty, and their primary responsibility is teaching. If a faculty member needs time to conduct a study, the institution must at a minimum relieve the individual from teaching. Ordinarily it will hire someone to teach in his or her place. Research grants provide the funding for the substitute instructor. If the researcher's employer does not like a proposal, it may decide against released time. A proposal that appeals to the views of learner-centered administrators and colleagues is more likely to find support.

Grants also pay what are called "indirect costs" for the use of the institution's facilities and other forms of overhead. These are additional funds that may amount to 50% or more of a research project's direct costs for a substitute instructor, equipment, supplies, and so forth. The funds an institution receives for such costs are typically added to various administrative budgets, thus enabling substantial discretionary spending. College

administrators consider faculty who generate big indirect cost contributions to be their most productive and deserving faculty. Grants are key to a faculty member's career advancement at major institutions. Grants that are readily funded for big amounts (e.g., grants from state education agencies) are thus extremely attractive.

Second, there is the matter of publication. In order to advance their academic reputations, researchers must publish. Research that is not published is assumed to be of lesser quality, and rightly so. Research that is published in the most respected journals is stringently peer reviewed. Reviewers and editors do not rule out findings that are inconsistent with orthodoxy, but such reports inevitably receive much closer scrutiny and are thus less likely to be accepted. A record of successful publication also contributes mightily to a researcher's chance of acquiring more funding.

Third, there is the matter of acceptance in the schools. The learner-centered view is more attractive to researchers because it is more easily marketed to the schools. Public school administrators typically have been trained in learner-centered thinking, thus such research has an intuitive appeal. That it may not produce intended results is a downside, but one that is frequently overlooked. School administrators are never fired or penalized because an innovative program fails. After all, how could an administrator be blamed for accepting the recommendations of scholar-experts who are supported by prestigious institutions. Because success is defined more in terms of funding than outcomes, appeal to decision makers is more important than demonstrated effectiveness. One need only observe the indicators of organizational advancement that are trumpeted in the media to verify the truth of this conclusion. Media releases talk about money and organizational expansion, not increased student learning.

The learner-centered view is comfortable to other stakeholders as well. Its convenience and vague expectations are significant considerations to teachers. In the learner-centered view, teachers are responsible for affording a quality educational experience, not the production of measurable academic outcomes. Learner-centered teachers consider outcomes to be governed by factors outside teacher control, thus the quality of teaching cannot be judged by results. Also, teachers find that learner-centered approaches are flexible and can be blended with existing practice without inconvenience and disruption. Factors of this sort make the task of adopting learner-centered practices simpler than, for example, implementing Direct Instruction--a methodology requiring more than the usual day or two of in-service training.

Learner-centered instruction also appeals to students. It seeks to accommodate them, not to shape them. By contrast, schooling that produces results typically requires a concerted student effort, and the time devoted

to such an effort can infringe on more attractive pursuits (Steinberg, 1996). It should be noted, however, that students' short-term satisfactions come at the cost of very substantial longer-term cost. Lost educational opportunity may result in permanently impaired career prospects--a delayed cost that students are unable to anticipate. Lost opportunities also cost taxpayers both in failed human resource development and the cost of remediation. Schooling that permits students to waste their own time and taxpayer-funded educational opportunity is an enormous but largely overlooked public disservice.

RECOGNIZING USEFUL RESEARCH

Research that can add to the efficiency and effectiveness of public schooling is available, but school personnel must be able to recognize it. Otherwise, there is a very substantial chance that they will be drawn into adopting one of the many fads that dominate the educational landscape. Recognizing credible, useful studies requires an understanding of certain basics of research.

Both medicine and education rely on a scientific knowledge base. Medicine, however, relies on relatively mature and exact sciences such as physics, chemistry, and biology, whereas education relies on the far less mature social and behavioral sciences. These differences in quality of research and precision of measurement are reflected in the certainty and internal coherence of the knowledge base on which the two professions rely. Competing and contradictory findings are not uncommon in the behavioral sciences; thus the matter of determining which findings are credible, important, and applicable is a formidable challenge to the educational practitioner.

Given facts open to selective use and interpretation, educators frequently rely on knowledge that is equivocal or that may be contradicted by other evidence. Recognizing this condition, Anderson, Reder, and Simon (1995) offer the following caution:

[N]ew "theories" of education are introduced into schools every day (without labeling them as experiments) on the basis of their philosophical or common sense plausibility but without genuine empirical support. [Instead] we should make a larger place for responsible experimentation that draws on the available knowledge. It deserves at least as large a place as we now provide for faddish, unsystematic and unassessed informal "experiments" or educational "reforms." We would advocate the creation of a "FEA" an analogy

to the FDA which would require well designed clinical trials for every educational "drug" that is introduced into the market place. (p. 24) Another limit on sound educational research is the inherent variability in human behavior. People think, feel, act, cooperate or don't cooperate, and so forth. Unlike inanimate objects, their actions are influenced by a range of extraneous variables that limit the applicability of findings. Behavioral sciences such as psychology have evolved standards that enable meaningful research despite these uncertainties. Unfortunately, many studies ignore them and consumers frequently fail to recognize the inevitable deficiencies and limitations. Thus it is not uncommon for educational administrators, grant writers, and program developers to stretch findings beyond their intended meaning or inadvertently to misrepresent results.

Quantitative versus Qualitative Research

Quantitative research includes both descriptive and explanatory studies. Descriptive studies are concerned only with establishing the existence of a phenomenon of interest--student achievement, for example. How much of it exists, where it exists, and what kinds of it exist are typical descriptive hypotheses. Explanatory studies are concerned with the causes of a phenomenon of interest. For example, does the use of Direct Instruction improve achievement? Technically stated, explanatory studies are concerned with the discovery of functional relationships (i.e., relationships in which the state of a given phenomenon is said to be a function of a preceding event or condition). Less technically said, explanatory studies are concerned with whether a given effect is the result of a particular cause. Causal relationships are examined in experiments and experimentlike studies called quasi-experiments. More is said about experiments below.

Descriptive studies address a wide range of topics. For example, a report of average test scores for students at different schools would be descriptive. So would a study of the number of words comprising recognition vocabulary of children at succeeding ages. Descriptive studies include a number of subtypes. For example, studies of characteristics such as preferred types of play or ability to perform certain intellectual tasks may entail observation of fresh samples of children at successive chronological age levels. Such studies are called "cross-sectional" descriptive research. Studies that examine the same characteristics but observe the same individual children over a period of years are called "longitudinal."

Quantitative descriptive studies also include reports of correlational relationships between variables. An example of a correlational study would be one that describes the degree of relationship between family socioeconomic status and school achievement. Another example is hyperactivity's relationship to junk food

consumption. Correlational studies are among those most frequently misinterpreted by users of educational research.

Despite its current unpopularity among educators, there is a great deal of high-quality quantitative research in education. It includes disquieting descriptive findings such as falling SAT scores and reports of low math and science achievement and similarly disquieting experimental results such as those of the Follow Through project. In the opinion of the authors, quantitative research's unpopularity may well be related to its disagreeable results. Findings that affirm orthodoxy are clearly more popular.

Qualitative research in education is a growth industry. It is a type of research long used in fields such as cultural anthropology. Qualitative research relies on written description instead of objective measurement, and its findings are subject to all the vagaries associated with written descriptions of any kind. Rather than attempting to affirm hypotheses and make generalizations that are grounded in an agreed-upon objective framework, qualitative research is more concerned with description as subjectively perceived by an observer in context. Such descriptions are thought to be more honest and realistic than descriptions that purport to be objective and at arm's length. It is a form of research premised on a postmodern, multiculturalist view of science. It argues that the objective understanding to which traditional science aspires is nothing more than an arbitrary Western convention--one educators should be free to reject.

By avoiding a focus on particular variables of interest, qualitative research presumably avoids the imposition of cultural bias. Of course such a process ignores the very information typically sought by the consumer. For example, a teacher's question about whether one teaching method produces greater achievement than another would not be answered by a qualitative study. Qualitative studies do not "prove" or "disprove" anything. They can only describe. The validity of such studies is simply an open question (Krathwohl, 1993).

The vagueness of the methods used in qualitative studies invites observer bias. Observers are necessarily selective in their observations. For example, an observer who dislikes the punishment seen in a classroom may tend to note the negative emotional reactions of students more than would a disinterested observer. By contrast, a more impartial observer might give greater attention to the increased on-task behavior that may be effected by the use of punishment. Although there are ways to make such observations more reliable, they are far more subject to researcher bias than most quantitative reports.

Action Research

Like qualitative research, action research has gained in popularity among educators. Wiersma (1995) describes it as research "conducted by teachers, administrators, or other educational professionals for solving a specific problem or for providing information for decision making at the local level" (p. 11). Action research is typically quantitative but less rigorous in design and methodology than conventional quantitative research. The following is a classroom level example: A teacher is having discipline problems during her fifth-period class. She arranges the desks differently and assesses whether the discipline problems are reduced. A written report of her investigation, including data, analysis, and a brief discussion, would be considered action research. Would such a finding be a sufficient basis for recommending that teachers employ rearranged desks as a means of treating discipline problems? In theory it would not. Practice, however, is another matter. Despite methodological weaknesses--in the present example, a single class sample and no control group--such findings are sometimes used to bolster proposals for new and innovative programs.

Pseudoresearch

Pseudoresearch is a form of scholarly writing that appears to make factual claims based on evidence but, in fact, consists only of opinion founded on opinion. Previous studies are cited, but they contain only theory and opinion. Legitimate empirical reports traditionally present a review of literature that enables the reader to put new findings in context and to strengthen factual generalizations (Stanovich,1996). However, previous studies containing only opinion do nothing to strengthen the report that cites them.

Commonsense educational claims are often supported by such "research." For example, if an expert opines that schooling is improved by greater funding and if other experts cite and endorse that original claim, subsequent reports will contain what appears to be substantiation. If the claim seems plausible and thus goes unquestioned, it appears to gain acceptance as a fact without ever being tested. Such claims are said to be supported by "research" but it is "research" in the sense of a systematic review of relevant literature, not in the sense of studies that offer an empirical foundation for factual assertions.

Educational innovations that are consistent with popular educational doctrines are often supported by such research. The controversial but widely used whole-language reading instruction (discussed below), for example, goes unquestioned by most educators because it fits hand-in-glove with learner-centered pedagogy. It is supported primarily by favorable opinion among like-minded educators, not demonstrated experimental results.

A type of research that seems to produce empirical facts from opinion is a group-interaction process called the Delphi method (Eason, 1992; Strauss & Zeigler, 1975). However, instead of creating the appearance of empirically grounded fact from multiple reports of opinion (as does pseudoresearch), the Delphi method creates facts about opinion.

In Delphi research, the opinions of experts are collected and synthesized in a multistage, iterative process. For example, if a researcher sought to determine the future occupations open to high school graduates, he or she might consult a panel consisting of career counselors, former high school students, employers, and economists. The panelists would be asked to compose a list of prospective jobs, and they would each share their list with the other panelists. After viewing the lists of other panelists some members might choose to change their estimations, and their changes would then be shared with the other panelists in a second round of mutual review. Ideally, three or so rounds of sharing and realignment would produce a consensus. The "fact" resulting from such a study is that experts agree about the future availability of certain jobs, not that certain jobs have a high probability of being available.

A recent attempt to find effective institution-to-home "transition strategies" for disabled juvenile delinquents illustrates how a Delphi consensus can be confused with an empirically grounded conclusion. Following three rounds of surveys, Pollard, Pollard, and Meers (1994) concluded that the priorities identified by the panelists provided a "blueprint for successful transition" when, in fact, the surveys produced only a consensus about what may or may not prove to be a successful blueprint.

Rand corporation is credited with developing the Delphi technique as a means of distilling a consensus of expert opinion. Sackman (1974) has summarized its primary shortcomings. The expert status of panelists is not scientifically verifiable and neither is the assumption that group opinion is superior to individual opinion.

One other confusion about the Delphi technique pertains to its use by the leader of a deliberative body. Delphi methodology can create the appearance of consensus where none exists--a problematic outcome of a deliberative process. Technically, the Delphi technique does not force a consensus; but as a practical matter, it is designed to produce a consensus and it puts substantial pressure on dissenters for conformity to the group. When employed by the leadership of a deliberative group, it can turn what should be an open and fair-minded exchange of views into a power struggle. Minority viewpoints can be isolated and marginalized. The result is more mindless conformity than reasoned agreement. The conclusions reached by committees and policy-making bodies can easily be distorted by Delphi methodology.

Experimental and Quasi-Experimental Research

Experiments are quantitative studies in which cause-effect relationships are tested (Campbell and Stanley, 1966). Quasi-experiments attempt the same but with certain limitations. Other studies may suggest or imply causal relationships, but their findings are far more ambiguous and subject to misinterpretation. Experiments are not foolproof, but they afford the best evidence science has to offer.

From a purely scientific standpoint experiments are important because they attempt to answer the primary question with which science is concerned: "What explains or accounts for the phenomenon under investigation?" All sciences aspire to this kind of understanding. They are valuable from a practical standpoint, too, because they address the question of whether a given program, teaching method, treatment, intervention, curriculum, and the like produces expected effects. Because schooling is intended as means of making a difference in the lives of students, the armamentarium of professional educators should contain tools that are well tested and demonstrably effective. Ideally, they should also be convenient, cost-effective, and well received by the student; but at a minimum, they must be effective. The critical importance of experimental evidence in establishing effectiveness is not well understood by educators, but it is just such an understanding that is at the heart of knowing which research is valuable and why.

The aim of science is said to be the explanation of natural phenomena. However, the term explanation itself requires a bit of explanation. As the term is used by scientists, explanation refers to cause-and-effect explanation. For example, a phenomenon such as achievement in school is said to be explained (or at least partially explained) if it can be shown that the presence or absence of achievement is functionally (i.e., causally) related to a preceding event or set of events termed a cause. A functional or causal relationship is initially stated in a tentative form called a hypothesis and is not considered a valid explanation until affirmed by evidence.

Experimental research is the business of collecting evidence that might support or disconfirm causal hypotheses. It entails the manipulation of a hypothesized cause for the purpose of inducing an expected effect. If a given effect (technically, a change in the "dependent variable") follows alteration of the purported cause (technically, a change in the "independent variable"), the causal hypothesis is said to be supported. Other types of quantitative research and even qualitative research may be valuable in suggesting cause-effect hypotheses, but only experimental research can provide a direct test.

Internal and External Validity of Studies

Whether an empirical study is capable of demonstrating a causal relationship is one issue, but whether a given experiment was properly conducted is another. Moreover, even a properly conducted experiment may have limited applicability and usefulness in the "real world." Whether the procedures used in an experiment permit valid findings is the matter of internal validity. Whether the findings of an experiment are generally applicable to the "real world" (i.e., applicable under conditions beyond those under which the study was conducted) is the matter of external validity.

A wide variety of technical considerations can adversely influence the internal validity of an experiment. For example, the manner in which subjects were assigned to treatment and comparison groups can profoundly affect the outcome of an otherwise well-designed experiment. Technical issues with respect to type of sampling and type of population sampled, for example, can greatly influence the external validity of a study. Accurate assessment of these and other technical details requires considerable expertise. Even well-informed investigators may overlook significant threats to the validity of an experiment. Cook and Campbell (1979) provide an authoritative discussion of the myriad considerations that should be considered. Happily there are at least three considerations that a nonexpert can examine to assess the internal validity of a study: source, convergence, and replication.

Source. If a study is reported in a peer-reviewed scholarly journal, chances are good that it meets acceptable standards of internal and external validity. Peer review typically entails blind review of a manuscript by a panel of experts selected by an editor. Panelists are not given the author's name and the author is not given the reviewers' names. All criticisms and replies are exchanged through the editor. The most reputable and selective journals use this process. Reports reviewed only by an editor may be valid, but peer-reviewed scholarship is generally conceded to be the most credible. Again, the process is not foolproof, but it is the best science has to offer. Unpublished reports and reports that are not subject to editorial review--grant proposals and reports of funded research such as those included in the ERIC's Research in Education, for example--are of uncertain quality and should be treated as such.

Convergence. If a study's findings are generally consistent with (i.e., they converge with) the findings of other investigations in an area of research, they are generally assumed credible (Stanovich, 1996). Any competent research report will include a review of relevant literature. Consistencies and discrepancies within the existing literature and between the report at hand and previous studies are analyzed and discussed. Articles called

"reviews of literature" and "meta-analyses" are dedicated to citing and summarizing all of the findings relevant to a given topic or area of study. Although new and revolutionary findings are sometimes uncovered by a single study, competent observations of the same or similar phenomena usually result in similar findings. Most scientific advancements come as incremental additions to understanding, not breakthroughs.

Replication. Replications are repeats of an original study by another investigator using a fresh set of subjects. The credibility of a study that has been replicated is greatly enhanced. Findings that have been replicated are considered valid even if they do not converge with other reports in the same general area of investigation. Only a small percentage of studies in the behavioral sciences are replicated, however.

The Need for Both Experiments and Field Testing

Few experimental investigations are able to fully satisfy requirements for both internal and external validity in a single study. The controls, artificial conditions, and other constraints necessary to ensure internal validity tend to interfere with external validity. Conversely, unanticipated and uncontrolled events can confound or invalidate an otherwise well-conceived study that is conducted in a natural environment such as a school. Because of this inherent conflict, programs or interventions derived from experimental investigations should be field tested prior to implementation.

Field tests are trials of an experimentally supported finding in the classroom or clinic or other setting for which it is intended. Not infrequently they result in the discovery of limitations, cautions, and restrictions on the applicability of experimentally validated findings. Even findings that have been field tested elsewhere may lack local applicability because of peculiar local conditions. Thus, large-scale programs, in particular, should also be locally tested on a small scale in what is called a pilot study. Pilot studies are especially important when the implementation of research findings entail significant time and energy costs for school personnel or learning opportunity costs for students.

RESEARCH AND EDUCATIONAL FADS

The failure of schools to employ reasonable precautions in adopting research-based innovations has been directly responsible for much wasted time, money, and educational opportunity (Carnine, 1993, 1995). Fads

are marketed like snake oil, and schools often adopt them with little credible evidence that they will work as promised. With taxpayers footing the bill and school personnel tracking the results, accountability is often minimal and adverse consequences rare. Typically once an investment is made, school personnel will either say good things about a program or say nothing at all. In the absence of a competitive marketplace, only students and taxpayers lose.

Despite the minimal risk entailed in adopting new programs, schools do seek to adopt those that seem most likely to succeed, yet they are often saddled with lemons. Naiveté about research is part of the problem, but so is ignorance of history. Today's fads are often nothing more than the latest incarnation of a philosophy or movement that has emerged and reemerged over many years. The learner-centered or child-centered or student-centered education concept noted below is the best example of an idea that has failed historically yet seems to have at least nine lives.

Still, one must ask why professional educators would be so gullible. Many are aware of history and most have the experience necessary to recognize what might work or not work. In a recent copy of Principal, a school administrator of 20 years offered the following reasons as to why educators are quick to jump on bandwagons:

• We believe and hear out of ignorance. • Most education programs for teachers and administrators do a pitiful job of teaching students to differentiate viable research from poor research. • We like doing whatever is in vogue. • We tend to move from one fad to another in order to demonstrate that we are "state of the art"even though most of the activities have little impact. • We seek a quick fix that will help all children succeed. If it doesn't work, we scrap it and try something else. But there are no quick fixes. • There is big money in selling educational programs. Consultants use "research says" to sell programs that purportedly can fix just about anything (Walker, 1996, p. 41). In the following section we discuss selected fads and innovations from the standpoint of their foundations in research. Clearly, considerations other than research play a role in their adoption, but our aim is to show how that in most instances a review and critical analysis of research would reveal weaknesses and raise questions about making a large-scale commitment of resources.

Current Fads

A prime illustration of how innovations with a weak or nonexistent research base can snowball into a movement is the current brain-based learning movement. The brain-based learning bandwagon has gained so

much momentum that it has been cited as grounds for federal legislative proposals. Briefer analyses of other recent fashions including constructivism, developmentally appropriate practice, situated learning, authentic assessment, learning styles, and whole-language follow.

Brain-Based Learning

Any educational proposition gains an aura of credibility if it can be tied to research in the hard sciences, and that is indeed the way in which brain-based learning makes use of neuroscience. The neuroscience on which brain-based learning is based is the product of legitimate research in the medical and biological sciences. Brain-based learning principles, however, are dubious interpretations of neuroscience, and their educational application is wholly untested (Bruer, 1997).

Making Connections: Teaching and the Human Brain, by Renate and Geoffrey Caine (1991), was published by the Association for Supervision and Curriculum Development (ASCD) and has been widely disseminated within the education community. Its recommendations regarding instruction are loosely based on "information from the neurosciences" (Regnier, 1996). The authors concede that direct translations of neuroscience findings to educational practice are risky and speculative, but they forge ahead with a list of suggestions.

For example, relying on Epstein's (1978) view that "spurts" in brain weight are related to increases in mental ability, Caine and Caine (1991) suggest that greater amounts of material can be taught during spurts and lesser amounts during suspected "slow growth" periods. Not only is their idea untested, it may have a more fundamental flaw: Epstein's idea is itself not well accepted by biological scientists; to the contrary, it has "long been known that there is little or no relationship between brain weight and brain functioning . . ." (Good & Brophy, 1986, p. 35). In other words, the educational effectiveness of their suggestion is not only unknown, it appears to be founded on questionable neuroscience.

Another claim made by both Caine and Caine (1991) and Healy (1990) is that neuroscience has demonstrated "brain plasticity" or an ability to adapt to new conditions throughout the life span. The notion of brain plasticity appeals to educators because it agrees with the popular educational concept of lifelong learning. Two problems are evident, however. First of all these authors use the concept differently from its use in neuroscience, and furthermore their idea is inconsistent with the evidence. Plasticity as the term is used in neuroscience refers to the ability of undamaged parts of the brain to take over the function of damaged

areas, not an ability to learn at any age. Second, neuroscientists have shown that true brain plasticity is greatest in young children, less in adolescence, and still less in adulthood (Pascual-Castroviejo, 1996). As do other proponents of brain-based learning, Caine and Caine appear to interpret and redefine neuroscience terms to suit their pedagogic purposes.

A third example of neuroscience interpreted in service of educational theory is Edelman's (1987) concept of "neural Darwinism." Edelman's view is founded on two analogies: that the brain can be thought of as a multilayered jungle and that it grows, changes, and adapts in much the same way as the immune system. Although Edelman's theory is only now undergoing study by neuroscientists, educational implications of neural Darwinism have been given cover-story treatment in widely circulated education periodicals. For example, in a 1994 Educational Leadership article, Sylwester asserts "Edelman's model suggests that a jungle-like brain might thrive best in a jungle-like classroom that includes many sensory, cultural, and problem layers that are closely related to the real-world environment" (1994, p. 50). Such a view fits nicely with Sylwester's apparent preference for unstructured, discovery-oriented pedagogy. What Sylwester fails to mention is that his interpretation is unsupported and his recommendation for classroom practice is in disagreement with a substantial body of evidence supporting the educational value of a well-ordered classroom.

Similar freewheeling interpretations of neuroscience are common. Cohen (1995) cites other proponents of brain-based learning as a basis for the assertion that educators need to throw out curriculum, textbooks, worksheets, and separate disciplines on the grounds that such curricular structure is inconsistent with our knowledge of how the brain works. Notably, he presents no evidence as to what happens with student achievement when such changes are implemented. Cohen goes on to say that many current "best practices" (i.e., portfolios, cooperative learning, and thematic curricula) are supported by brain research yet cites no research to vindicate these claims. Plainly, Cohen's references to neuroscience are nothing more than rhetorical props for his beliefs about "best practices" in education.

The value of assessing research-based claims on the basis of their source is well illustrated in the case of Sylwester's "What the Biology of the Brain Tells Us about Learning" (1994). The majority of Sylwester's references were published by one sourceBasic Books. Basic Books, a commercial publisher, is also the publisher of Edelman's popular account, Neural Darwinism (1987). Not only were they mostly from one commercial publisher, Sylwester's scientific references were drawn from one source--a special issue of Scientific American. Well-founded claims typically have a much broader base in research, especially claims that are sufficiently well confirmed to serve as a guide to classroom practice.

Jane Healy is another widely known proponent of brain-based learning. Her Endangered Minds: Why Children Don't Think and What We Can Do About It (1990) suggests that societal changes have caused changes in brain structure that are responsible for deficient student achievement. She brings together many recent neuroscience findings, but like other proponents of brain-based learning she selectively draws implications that serve to support "pedagogically correct" views of teaching. For example, Healy says that "research has shown that good readers actively pursue meaning . . ." (p. 298), but offers no research in support of her claim. Her view is really nothing more than an attempt to lend credibility to the widely accepted "constructivist" view of reading.

Like Caine and Caine's Making Connections (1991), Healy's book is published by a nonacademic publishing house. It was intended primarily as an explanation for why kids today seem so different from those of previous generations. Many of Healy's ideas are intuitively appealing (e.g., we need good teachers; tailoring the school day and school year to families' schedules will help students and parents; children should be taught to listen effectively) but they are not supported by research. This is not to say that they are demonstrably wrong; rather, they are speculative and unsupported by credible evidence. The same can be said about her hypothesis that society's fast pace and electronic complexity have caused fundamental changes in children's brains. At best, the evidence is merely suggestive and subject to other interpretations. Deficient evidence notwithstanding, Healy draws conclusions and pronounces learner-centered educational orthodoxy vindicated. Her recommendation: If brains have been adversely affected by the environment, schools must change to accommodate them--whatever impact such changes might have on the outcomes of schooling. The idea that schools might act to shape student thinking processes in a way that is conducive to the intellectual characteristics commonly associated with a good education is not considered.

Much more can be said about educational claims that derive from neuroscience, but all suffer from the same general flaw. Bruer states it simply: "Currently, we do not know enough about brain development and neural function to link that understanding directly, in any meaningful and defensible way to instruction and educational practice. [Furthermore] we may never know enough to be able to do that" (1997, p. 4). In truth, brain-based learning appears to be little more than one more attempt to justify the learner-centered educational doctrines that have dominated the education community for decades.

Constructivism and Developmentally Appropriate Practice

Constructivism is an educational doctrine founded on the idea that each individual constructs his or her

own understanding and knowledge from personal experience. It implies that schooling should concern itself not with the acquisition of an accepted body of knowledge but with the process of helping students discover and create their own understandings. It is consistent with the poststructuralist and deconstructionist perspectives in literary theory. In education, it is consistent with the progressive/learner-centered view of learning articulated by John Dewey. Dewey held that education should result in an intellectual "growth," not the achievement of preconceived educational outcomes.

Present-day educational constructivism is primarily tied to the concept of intellectual development formulated by the Swiss theorist Jean Piaget. Like Dewey, Piaget viewed intellectual growth as the prime outcome of education and experience the best teacher. Piaget's concept of "adaptation" argues that children construct a personalized grasp of the world by alternately "assimilating" various understandings of the world (called schemata) and refining those understandings through "accommodation." The aim of schooling from the Piagetian perspective is to optimize the "growth" or "adaptation" of the individual by fitting educational experience to the characteristics and proclivities of the individual student. Attainment of conventionally measured student achievement is a secondary and incidental outcome (Stone, 1996).

Piagetian constructivism is the theoretical foundation for what the National Association for the Education of Young Children (NAEYC, 1991) calls "developmentally appropriate practice" (DAP). DAP seeks to facilitate the construction of understanding (i.e., intellectual development) in ways that are compatible with level and pace of the individual's developmental trajectory. It is thoroughly child centered in the sense that children are not prodded or induced to undergo experiences that might be incompatible with what Piagetians suppose is a naturally shaped and therefore optimal developmental progression. DAP avoids subjecting the child to any sort of normative expectations for effort or accomplishment because even these subtle pressures might put a child's longer-term intellectual development at risk.

The practical problem of knowing a student's current level of intellectual development places an additional restriction on developmentally appropriate teaching. Because intellectual development can only be known through overt performance and overt performance is influenced by both learning and maturation, a child sufficiently mature to engage in intellectual tasks beyond his or her present level of performance may not appear ready for instruction. In effect, DAP encourages teachers to await the appearance of intellectual readiness even if a child's apparent lack of readiness is due to deficient motivation--a waiting period that may place the individual far behind peers (Johnson & Johnson, 1992).

DAP reduces a hypothetical risk to intellectual development, but it does so at the expense of teachers' taking a very passive role in fostering academic attainment. However, from the standpoint of DAP's proponents and that of other constructivists, the delayed academic progress of some students is not any legitimate grounds for criticism. In their view, DAP is intended to produce a pattern of intellectual growth unique to the individual, not a pattern of achievement that compares favorably to norms. Thus, exponents would reject the view that DAP is ineffective merely because students fail to learn as defined by conventional measures. Rather they believe that DAP protects children from overly ambitious expectations--a questionable tradeoff in the view of the few parents and other consumers who understand DAP's aims.

Despite DAP's rejection of age-referenced academic expectations, the effectiveness of teaching practices very similar to DAP"open education" and "discovery learning"were investigated in the 1960s and 1970s. Open classrooms sought to take away the desks in rows and teacher-directed classroom activities and replace them with rooms containing learning centers and student-directed exploration and discovery. In truth, these innovations of the 1960s were reincarnations of the child-centered classrooms of the 1930s. Student freedom afforded in a facilitative school environment was expected to incite "intrinsic motivation." Learning that had formerly required orderly didactic instruction was expected to emerge as a result of a spontaneous engagement with interesting activities and materials.

The experiment failed. Discovery learning and other experiential methods were found to be more expensive and more time consuming, and they left behind many students who just did not seem to blossom despite facilitating conditions (Good & Brophy, 1986; Rosenshine, 1978). With regard to discovery learning, "the larger, better controlled studies tend[ed] to favor traditional education--especially on achievement measures" (Good & Brophy, 1986, p. 212). Open education was a particularly visible "bust." In a meta-analysis of 153 studies, Giacomia and Hedges (1982) found small negative effects in all areas of achievement and a substantially greater deficit in achievement motivation. In other words, students experiencing the DAP style of teaching not only failed to learn but failed to acquire the motivation necessary to subsequent school success.

John Anderson, Lynne Reder, and Herbert Simon at Carnegie Mellon are among the foremost cognitive psychologists in the United States. Their take on both constructivism and situated learning (discussed below) is that the claims are unproven and, in several respects, at odds with well-known scientific findings (Anderson, Reder, & Simon, 1995). Moreover, their practical worth in the classroom is suspect at best. As with so many other widely known innovations, they flourish within the education community not because they are

supported by sound research but because they are well accommodated to the prevailing learner-centered orthodoxy and unchallenged by demands for accountability.

Situated Learning

Situated learning is another example of an idea about teaching that has gained acceptance within the education community not because of its demonstrated effectiveness but because of its compatibility with the learner-centered perspective. Again, educational concepts that are very like what is now called situated learning can be found in writings of John Dewey and William Heard Kilpatrick. Both called for a curriculum built around student projects and both believed that experience is the best teacher. In Dewey's view, human evolution had selected intellectual abilities that made learning from experience the most natural and most effective form of education (Stone, 1996). Dewey argued that challenging students with true-to-life problems would eliminate the rigors and artificialities of traditional classroom instruction while making schooling both more effective and more attractive to the learner.

The core principle of situated learning similarly argues that anything learned in the context of the situation to which it will be applied is learned more thoroughly and in a more useable form. Situated learning is thought to provide not only context for learning but an incentive that is missing from classroom exercises. Situated learning is consistent with the idea of "hands-on" educational experiences and it argues for the use of "authentic assessment" (discussed below) rather than conventional tests and classroom exercises. For example, the situated-learning approach to teaching math might entail student participation in a construction project. Otherwise boring math skills could be practiced by estimating the quantity of needed materials. Presumably the project would make the learning activity more attractive (i.e., intrinsically interesting and socially engaging) and the circumstance would ensure that the acquired math skills would be integrated with other skills such as measurement and thus made useful in the real world. Isolated and decontextualized "book learning" would be avoided. Too, failure or success in completing the project would give students real-world feedback about the quality of their skills as well as provide the teacher a visible indication of how much and how well they had learned.

Despite its motivational appeal and the merits of learning in context, enthusiasts give little attention to situated learning's prospective shortcomings. Its primary weaknesses are inefficiency, cost, and uncertainty of outcome. Traditional schooling attempts to teach by breaking the learner's task into a series of simpler, more manageable tasks and building on these basics. Situated learning starts learners at the application level

and attempts to teach the basics later. It appeals to students, but its effectiveness has not been demonstrated. Individuals who do learn via such experiences may be highly motivated at the outset, but reality can quickly take its toll. For example, novice tennis players who begin by attempting to play competitively quickly discover that stroking the ball is not as easy as it looks on television. Frustration and discouragement theoretically are overcome by mentoring, but whether they do so in fact has yet to be empirically validated.

In the judgment of Anderson, Reder, and Simon (1996), the claims made for situated learning are not only excessive, they ignore or reject much that is known about the value of abstract and decontextualized learning. Situated learning is recommended when time-tested practices might be simpler and more effective. "It is a well documented fact of human cognition that large tasks decompose into nearly independent subtasks" thus enabling simplification of the learner's task (Anderson, Reder, and Simon, 1995, p. 3). For example, the acquisition of useful computation skills does not require one to learn addition and subtraction in the context of doing one's income tax. Moreover, in at least some areas of instruction, abstract and generalized learning is more efficient than contextualized methods because it only requires the learner to apply known concepts to new circumstances. By contrast, learning closely tied to a specific task and context may require the learner to retrain completely for each new application.

Other researchers argue that situated learning can be detrimental. Shiffrin and Schneider (1977) and Winn (1994) suggest that specific skills learned in context may actually work to the learner's disadvantage. They conclude that the automaticity associated with expert performance of a given task may encourage the development of inflexible and difficult-to-generalize skills. In other words, at least some skills are more useful if they are taught as abstractions rather than in context.

Anderson, Reder, and Simon (1996) object to other situated learning's claims as well. Citing a body of psychological research that has accumulated since the mid-1800s, they argue that situated learning's proponents ignore a large number of studies that show that training in one context can be transferred to a novel situation or task. For example, skill in using one word-processing editor can make the acquisition of proficiency in the use of a second editor far quicker and easier. Thus contrary to the studies that have been highlighted by situated-learning proponents--ones in which transfer did not take place (Glick & Holyoak, 1980, 1983)Anderson, Reder, and Simon find successful transfer a common occurrence. Not only is transfer common, it can often be prompted by something as simple as a suggestion that a previous task and a new task have certain features in common. Again, their conclusion is that classroom exercises and "book learning" do not somehow disadvantage the learner (Hunter & Hunter, 1984) but, in fact, are often enabling.

Situated learning and the many other variants of contextualized learning have a huge intuitive appeal. Everyone can think of personal examples in which they learned well through experience and mentoring. The problem is that much of what students are expected to learn in schools requires understanding of the decontextualized symbols and abstractions that represent the distilled experiences of previous generations. Unassisted access to this wealth of experience requires proficiency in accurately decoding and making use of information represented in this form.

The primary adaptive value of education is that it equips the learner for future conduct. Learning in a naturally occurring context may be a richer and more meaningful experience, but the advantage of the educated over those familiar only with immediate experience is that the educated have much greater knowledge of and access to the experience of others. In the face of uncertainty, the educated are better equipped to anticipate which paths of action lead to which ends. Thus possession of a fund of abstract and symbolic information and demonstrated proficiency in decoding the enormous body of written and spoken knowledge is a significant adaptive advantage that seems slighted by situated learning.

Dewey may have been correct in arguing that the human species is better equipped to learn from experience than from exposure to symbols and abstractions. In fact, his view is consistent with the observation that the human ability to compress past experience into symbolic information and to store it outside the body is a relatively recent development in human history. Yet it is precisely because the benefits of widespread access to such information are so enormous that all modern societies have some system of formal education; and in an information age, proficiency with symbols and abstractions are of unprecedented importance.

Even if decontextualized, learning is more challenging and less immediately satisfying than hands-on, contextualized experience, and competence with symbols and abstractions must be education's chief priority. Situated learning and other forms of learning in context may be attractive to the learner, but their cost effectiveness in producing adequate levels of literacy and numeracy are unproven at best. An individual's opportunity to learn is not unlimited. Students have only so many unencumbered years. Adults have only so many years during which they can make use of that which they have learned. Socioeconomic communities compete with each other, and investment in education pays off only to the point that it remains an economic advantage. Schooling that places student satisfaction ahead of these economic realities is disadvantageous and in the longer term unsustainable.

Authentic Assessment

Authentic assessment is a means of measuring student learning in which lifelike tasks or their products are observed. For example, the ability to use nouns and verbs correctly might be assessed by observation of students writing a letter rather than by counting correct responses to exam items. The ability to read might be assessed by observation of students using a recipe to bake a cake rather than by listening to them reading. Other names for this kind of assessment are direct assessment and performance assessment. The term portfolio assessment refers to the practice of collecting "best" authentic products over a period of time as a measure of overall achievement.

As is the case with so many educational innovations, authentic assessment is well regarded less because of its value in precisely appraising academic progress than its compatibility with a learner-centered vision of schooling. It fits Dewey's view that the outcomes of schooling should reflect the aims of the learner, not the aims of some external agent such as a school board. It is compatible with constructivism and situated learning and for the same reason. Each viewpoint reaches the learner-centered conclusion that standards for learning should consider that which is important to the learner rather than merely that which is important from the standpoint of an imposed standard.

There are shades of opinion among authentic assessment proponents as to the precise role of external standards in judging quality. At the extreme are constructivists who argue against the application of any standards. In the opinion of Anderson, Reder, and Simon (1995, p. 20), "The denial of the possibility of objective evaluation could be the most radical and far-reaching of the constructivist claims." For example, Madaus (1994) argues against tests on the grounds that they construct, control, and dominate social persons and thus are instruments of social and political control.

The primary argument on behalf of authentic assessment is that it avoids student learning intended merely to pass tests rather than gain integrated knowledge and useful skills. It does so by requiring application-level educational outcomes in which the individual learner presumably has a stake. Its weakness, however, is that application-level performances can be superficial and misleading, and assessments based on them can overlook important aspects of learning (Baker, O'Neil, & Linn, 1993).

Application-level competence is clearly a desirable schooling outcome, but it is assumed that such demonstrations are an expression of more generalized understanding and ability. In fact, successful

performance under authentic conditions may or may not represent a grasp of critically important knowledge and skills. For example, if students working in a cooperative group successfully repair an automobile, their performance may seem to demonstrate that they are able to read a repair manual and order parts correctly. In truth, it may indicate only that they are able to follow the advice of a knowledgeable friend or parts store clerk.

The key difference between authentic assessment and conventional educational measurement lies in the purposes they are best suited to serve. Conventional assessments measure whether students are able to decode accurately and interpret information presented in symbolic and abstract form. Can they accurately interpret a communication? Do they get the message? Do they understand the question and know the relevant information? Authentic assessment is best suited to determining whether students are able to carry out a task or activity successfully. Can they solve the problem? Can they complete the task however they go about it?

Performance founded only on previous experience in doing an activity or task may tell the observer nothing about the learners' knowledge and understanding or their ability to make use of knowledge and understanding. Exclusive use of authentic assessment not only makes possible schooling that ignores learner capability with the use of symbols and abstractions, it encourages the exclusive use of the hands-on, experience-oriented teaching methods that may avoid reading and writing entirely.

Assessment of student ability to engage in real-world tasks and activities has traditionally been a part of classroom instruction. What is new with the current emphasis on "authentic assessment" is the notion that such assessments should be used exclusively (i.e., in place of conventional tests, even where the assessment is a basis for educational accountability). If schools are to be accountable for both conventional and application-level outcomes, both types of assessment are needed.

Authentic assessment is plagued with other difficulties as well. From a technical standpoint, there are problems with reliability, validity, and cost (Willson, 1991). Reliability is the matter of consistency in measurement. For example, if the readings of a bathroom scale vary only slightly when an individual steps off and on again, its readings would be termed reliable. Of course, it could be consistently wrong by 50 pounds. If, in addition, the scale's readings correctly showed the weight of a 100-pound object to be 100 pounds, the scale would also be called valid.

Because reliability is a prerequisite to validity, authentic assessment has problems in both areas. Authentic assessment typically requires observations of products and performances made by multiple observers. Often the observers disagree by unacceptably large amounts. If observers cannot agree as to the quality of a product or performance, their reports--individually or collectively--cannot be treated as valid.

Despite strenuous efforts to improve consistency through the use of trained raters, detailed scoring guidelines, and other means, inter-rater reliability remains an expensive problem and major limitation. Hoover (as cited by Willson, 1991) estimates the extra expense entailed by authentic assessment to be 10 to 100 times that associated with traditional methods of testing. In contrast, less expensive traditional tests "(a) are actually effective; (b) are free of unwanted negative consequences; (c) meet established and reasonable psychometric criteria for validity, reliability and freedom from bias" (Hambleton & Murphy, 1992, p. 10). Hambleton and Murphy modestly suggest that much more research needs to be done before authentic assessment is used exclusively.

Authentic assessment's popularity is not founded on a body of findings that show it to be a vast improvement over conventional testing. If anything, it is less reliable, less valid, and more expensive (Baker, O'Neill, & Linn, 1993). Rather, it has gained the attention of educators because it sets real-life performance as the defining measure of school achievement. From the standpoint of learner-centered pedagogy, such indicators are an enormous advancement over the narrow, fragmented, and decontextualized knowledge and skills required by traditional objective tests. However, as a means of assessing student understanding and proficiency in the use of symbols and abstractions--the core of what it means to be an educated person--authentic assessment is uncertain at best.

Cooperative Learning

Cooperative learning is another "hot" educational methodology. Woolfolk (1998) defines it as an "arrangement in which students work in mixed-ability groups and are responsible for each other's learning" (p. 350). The presumed advantage of cooperative learning is that students will gain in cooperative social skills while achieving academically at levels equal to or greater than those associated with traditional instruction. Instead of encouraging students to compete, cooperative learning encourages them to assist each other. Instead of encouraging students to rely on teacher direction, cooperative learning encourages them to rely on each other.

Cooperative learning is only the most recent incarnation of an idea that has emerged repeatedly in American educational history: instruction through small-group interaction (Lorge, Fox, Davitz, & Brenner, 1958). Eighty years ago, William Heard Kilpatrick's The Project Method urged schooling centered around group projects. Very similar methodology is today called "project-based learning" (Stern, 1996) and is thought to be the cutting edge of educational innovation. Another version of learning via group interaction is the discussion-group methodology that was popular in the 1950s and 1960s (Hare, 1962). Both project groups and discussion groups were attempts to harness group-interaction processes to the task of teaching. Both methods improved student motivation and student satisfaction, but neither succeeded in significantly improving conventionally measured achievement. Again, as with so many teaching practices, project-oriented groups and discussion-group methods have gained popularity not because of their demonstrated effectiveness but because of their fit with the learner-centered view of teaching. From the learner-centered perspective, if man is by nature social and knowledge is a social and cultural construct, then an interactive social context must be more naturally conducive to learning than the traditional isolated learning activities.

Does cooperative learning work as advertised? The answer depends on which of its many versions is considered and for which students. There is a rather large body of research on cooperative learning, and overall it shows that cooperative learning does produce modest achievement effects (Slavin, 1995). A possible weakness, however, is that not all members of cooperative groups benefit, and it may be that cooperative learning benefits lower-achieving students at the expense of those who are more talented (Druckman & Bjork, 1994). Cooperative learning groups are typically heterogeneous, meaning they are comprised of students with varying ability levels. Stevens and Slavin (1995) found effective benefits and achievement gains for all students as a result of cooperative practices in an elementary school. Special education students, however, made the greatest gains. A study reported by Bramlett (1994) similarly found achievement gains only for the lowest one-third of students. No studies of cooperative learning have found exceptional benefits for high ability students.

Cooperative learning has gained a reputation as an innovative teaching methodology that is more than a mere fad. It works. Robert Slavin (1995)the creator of cooperative learning identifies two crucial elements to the successful use of cooperative methods: There must be some type of reward or recognition for the group and there must be individual accountability for members of the group. When these two conditions are met, cooperative learning usually succeeds. However, should cooperative learning be considered a primary tool for teaching in the public schools? The answer is a matter of educational priorities.

Clearly, cooperative learning does produce achievement and it is an improvement over various other forms of instruction, but should schools settle for modest achievement gains or limited gains with more talented students as a price of achieving social and emotional ones? In other words, should social and motivational outcomes be put on an equal plane with academic outcomes? The teaching profession may say yes, but the public would probably disagree. On balance, parents and policy makers want achievement to be an unrivaled priority. Most parents, especially parents of intellectually talented students, want their child's abilities maximized, not constrained by socially oriented pedagogy.

Learning Styles, Individual Differences, Diversity, and Attribute-Treatment Interaction Research

The idea of fitting instruction to the unique characteristics of the student is one of the most intuitively appealing notions in pedagogic theory and one of the oldest modes of learner-centered education. It originated with the child study movement of the early 1900s (Spaulding, 1903) and has been researched repeatedly since (e.g., Davis, 1948; Caplan & Ruble, 1964). In the view of proponents, if learners have unique social, emotional, and intellectual characteristics, it should be possible to optimize learning by fitting schooling to them. Not only has learning style research been dedicated to this idea, many studies that are technically concerned with aptitude, personality, and developmental assessment have been enlisted in the effort. These areas of investigation are not all recognized as such but, in principle, all might be included under the broad heading of attribute-treatment interaction research (Snow & Swanson, 1992).

The general problem with regard to studies of student attributes and their relationship to achievement is that numerous differences in personality, intelligence, learning styles, and other characteristics have been described, but appropriate intervention has been only suggested or left to the imagination of the teacher. In other words, they are long on diagnosis but short on treatment. The many studies identifying learning styles illustrate the problem (Entwistle, 1981; Snow, 1992). As Slavin puts it: "What has never been studied, to my knowledge, is the question of whether teachers who adapt to students' styles get better results than those who don't" (cited in Ellis & Fouts, 1993, p. 69). Studies of multicultural diversity are broad attempts to identify race, gender, language, and other group differences that correlate with educational outcomes--again, for the purpose of better fitting schooling to the student. They too are long on diagnosis and short on proven interventions. Many differences are identified, but the matter of how to fit schooling to those differences in

some advantageous way is left unanswered. In truth, the attention given studies of individual differences and diversity bears little relationship to their usefulness in improving academic achievement. Rather, they are a product of Dewey's learner-centered view buttressed by social and political considerations.

The core problem in interpreting research on student differences is knowing whether the relationships that have been discovered are functional relationships (i.e., causal relationships) or merely incidental correlations. In other words, do the correlations between race or gender or learning style and school success mean that schools can take some action that would improve outcomes? Despite any clear indication of what, if anything, about the school environment may be responsible for some groups performing less well than others, schools are frequently stampeded into making changes and accommodations that generally presume that diversity has been insufficiently accommodated or welcomed. Changes in teaching, organization, funding, hiring practices, curricular content, faculty training, pupil assignments, and leadership are only some of the responses that have been undertaken and, in general, they have shown little systematic relationship to achievement.

A case can be made that accommodations to race and gender differences, for example, have at least resulted in gains such as lessened stereotyping and greater racial tolerance, but even these nonacademic gains are not an unmixed blessing. Interventions based on a premise that differences are related to deficiencies suggest the condescending notion that differences are disadvantages--a view that contains the seeds of self-fulfilling prophecy and a ready-made excuse for failure.

Whole-Language

Whole-language reading instruction is premised on the idea that children can and should learn to read "naturally" (i.e., through the same socialization processes that teach them to speak). It calls for reading instruction to be indirect, unsystematic, and nonintensive and it assumes that structured, sequenced, skill-building approaches to reading are likely to be harmful. Again, the influence of learner-centered education is evident.

Everyday observations seem to contradict whole-language's assumptions. As the Canadian Organization for Quality Education has observed, virtually every toddler learns to speak and communicate, but there is a large number of adults who never learn to read (Dare, 1997). In a world in which the printed word is ubiquitous and literacy is common, how could so many not learn? Plainly, even if there is some natural socialization

process that will produce competence in reading, the necessary conditions are not as widely available as those that engender speech. Repeated studies of whole-language versus phonics-based reading instruction have found the structured phonics-based methods to be more effective--especially with at-risk students (Chall, 1967/1983). Yet, disagreement persists about the quality of the underlying research and even which outcomes are most important (Carbo, 1988; Turner, 1989). A recent summary of experimental findings, however, makes it clear that if schools are to reduce the growing number of poor readers--40% of third grade students lack proficiency--they need to think of phonics, not whole language (Grossen, 1997): "Treatment intervention research has shown that appropriate early direct instruction seems to be the best medicine for reading problems. Reading is not developmental or natural, but is learned. Reading difficulties reflect a persistent deficit, rather than a developmental lag in linguistic (phonological) skills and basic reading skills. Children who fall behind at an early age (K and grade 1) fall further and further behind over time."

Previous comparisons have found little support for whole language. Stahl and Miller's (1989) comprehensive review found no evidence that whole language produced stronger effects than basal programs, and in a number of investigations, poorer results were reported. Examining the conceptual basis for whole-language, Vellutino (1991) states, "I think it is fair to say that the major theoretical assumptions on which whole-language approach to instruction are based have simply not been verified in relevant research testing those assumptions" (p. 442).

By contrast, there is a wealth of experimental research demonstrating the value of reading instruction using artificially taught decoding skills (i.e., phonics). Foorman (1994) compared phonics-based reading instruction with whole-language and found phonics clearly superior. Her findings are consistent with those of Paulu (1988), Adams (1990), Brown and Felton (1990), Engelmann (1992), Groff (1994), and Sears and Keogh (1993), among others. The accumulated evidence seems clear as to which methodology is best suited to producing the educational outcomes wanted by parents and policy makers.

Current Fads and Innovations: A Common Impediment

The learner-centered vision of schooling is a thread to which all of the above innovations are linked, and it contains a concept that is responsible for both their marginal effectiveness and their lack of experimental support: It is the idea that true learning can come only from an inner or intrinsic motivation. Learner-centered schooling presumes that a well-fitted learning environment will produce a spontaneous and "intrinsically motivated" student effort--one commensurate with optimal learning outcomes for the student

(Stone, 1996). If the student fails to experience such an urge, the deficiency is presumed due to some lack of environmental accommodation, not a shortcoming on the part of the student. In other words, the student is expected to act only out of a genuine sense of interest and enthusiasm, not one of responsibility or dedication, and certainly not in response to any extrinsic pressures or inducements.

Such a perspective presents a formidable if not insurmountable challenge to the teacher. Extensive empirical evidence demonstrates that learning requires time, attention, and action on the part of the student (Pirolli & Anderson, 1985; Steinberg, 1996), yet the learner-centered teacher is permitted only to accommodate the student's needs and to otherwise await the student's cooperation. Whether a teacher views his or her task as one of accommodating to spurts in brain growth or engaging students in cooperative activities, student effort is needed. Teacher-directed intervention, however, is considered an intrusion into naturally occurring learning and development processes and is presumed capable of causing harm.

In the end, the learner-centered vision leaves teachers and schools with the responsibility of getting students to undertake the activities that are essential to educational success, but they are sharply restricted in the kind of steps they can take. They are expected to energize and excite but prohibited from directing, inducing, impelling, or otherwise taking action that might not be well received by the student. In other words, from the student's standpoint, effort in the learner-centered school is optional and expected only if the individual "feels" so inclined.

The practical teaching and learning problems that flow from the learner-centered view are obvious to experienced educators, yet the learner-centered orthodoxy is so powerful that more directive teaching practices are considered beyond the realm of competent and ethical discussion. Many who administer America's public schools believe that students are harmed if teachers insist that they pay attention, study, and behave themselves. In the absence of adult insistence on appropriate constructive activity, students are drawn to the activities encouraged by peer groups, commercial interests, and other noneducational influences.

A second shortcoming of the above-discussed innovations stems from the way in which the learner-centered vision frames the question of teaching effectiveness. Again, the root difficulty is the idea that true learning is a function of intrinsic or endogenous factors, not some external agent. School and classroom conditions are thought to influence outcomes only to the extent that they accommodate the unique proclivities of the individual learner. These external or exogenous conditions are considered relevant not as agents of change but merely as conditions that facilitate the working of endogenous agents. In other words, the actions of

teachers and schools are not intended as independent variables that stand in some causal relationship to schooling outcomes. Rather, they are ad hoc adjustments made in anticipation of or in response to student idiosyncrasies that presumably permit the successful working of the endogenous causes of learning.

As an example, the developmentally appropriate practice viewpoint holds that learning occurs when development has advanced sufficiently and appropriate facilitating conditions are in place. The factor primarily responsible for learning is the student's endogenous development, and the teacher's contribution is a secondary one of affording proper accommodations. Conceptually, the teacher's actions do not produce the outcome. Rather, the teacher facilitates an outcome that is some unique product that issues from the confluence of developmental processes and exogenous conditions comprising school, home, and all other influences. Constructivists, for example, say that the resulting student understanding is "constructed" by the student. Thus, experiments in which teacher actions and school conditions are treated as causes of learning are conceptually questionable.

Given such a framework, it is not surprising that most learner-centered educational practices have only been the subject of descriptive studies, not experiments. The actions of the endogenous developmental agents can only be inferred from secondary indicators such as age or student success in performing certain tasks. Measurement of these indicators permits researchers to conduct studies in which academic achievement is forecasted (i.e., descriptive or correlational studies), but the less ambiguous experimental studies are largely precluded. Experiments require systematic manipulation of "independent variables" (i.e., hypothesized causal influences) for the purpose of affirming or disconfirming their impact on outcomes.

Effective, Research-Based Educational Methods

There are studies that demonstrate effective teaching methods, but they are not well accepted by a teaching community that has been indoctrinated in the learner-centered thinking. By and large, the effective methods conceive of teachers as causal agents rather than facilitators of learning induced by other agents. References to many of these reports can be found in the articles noted above by Ellson (1986) and Walberg (1990, 1992). Other reviews of literature and meta-analyses of research add to the conclusion that effective methodologies are available but little used (Ellis & Fouts, 1993; Lipsey & Wilson, 1993). In most cases, the effective interventions are structured, teacher-directed, and designed to produce preordained academic and intellectual outcomes. In contrast to the interventions such as brain-based learning, developmentally appropriate practice, and others, most are learning-centered, not learner-centered (i.e., they are intended to induce or impel certain

activities and outcomes, not merely energize and excite).

Broadly speaking, these are studies of methodologies intended to produce what the public wants from its schools: measurable academic achievement. The mainstream education community ignores or rejects them as antiquated, too artificial and mechanical, overly reliant on extrinsic motivation, and generally unenlightened. The evidence, however, generally illustrates that they are more effective and efficient than the prevailing learner-centered approaches.

WHICH RESEARCH, WHICH AIMS?

This chapter began with the question of how there could be so much research yet so little improvement in the achievement of students, and it has attempted to show that much of popular educational research is not intended to serve the public's educational aims. Choosing credible research is to a very great extent a matter of understanding the educational aims that a particular piece of research is intended to serve. There is a body of educational research that in a reasonably impartial way demonstrates the effectiveness of various educational methodologies and the ineffectiveness of others. Examples of such research have been cited throughout this chapter. They include the Follow Through findings (Proper & St. Pierre, 1980; Engelmann, Becker, Carnine, & Gersten,1988) and the many studies cited by Ellson (1986), Walberg (1990, 1992), Lipsey and Wilson (1993), and others. Such reports are plentiful and found in journals such as Review of Educational Research, Journal of Applied Behavior Analysis, the American Psychologist, and numerous other peer-reviewed scholarly journals.

There is also a substantial body of educational research that is said to represent the best and most up-to-date findings, but it is dedicated to conceptions of education that do not consider academic achievement to be the highest priority (see Heffernan, 1958, for an excellent illustration of turn-of-the-century concepts introduced as the cutting edge of educational thinking). Many of these studies have also been cited above. It is research that places objectives such as student satisfaction, enhanced self-esteem, equity, social justice, and other nonacademic outcomes as equal or superior to academic achievement. It is very often nonexperimental or even nonempirical, and its objectives, however laudable, are not advantageous from the standpoint of optimized academic achievement. Frequently it is this type of research that is cited in support of popular educational innovations. Educational innovations based on such research are adopted not because they afford

the kind of education most parents want for their children but because they constitute what many educators believe to be a superior view of schooling (i.e., a learner-centered view). Thus there is much research and relatively little improvement because the kind of research that best suits the public's aims is only infrequently implemented by the schools.

References

Adams, M.J. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Anderson, J.R., Reder, L.M., & Simon, H.A. (1995). Applications and misapplications of cognitive psychology to mathematics education. Unpublished manuscript. Department of Psychology, Carnegie Mellon University (accessible at http://www.psy.cmu.edu/~mm4b/misapplied.html).

Anderson, J.R., Reder, L.M., & Simon, H.A. (1996). Situated learning and education. Educational Researcher, 25(4), 5-11.

Baker, E.L., O'Neil, H.F., & Linn, R.L. (1993). Policy and validity prospects for performance-based assessment. American Psychologist, 48(12), 1210-1218.

Bramlett, R.K. (1994). Scientific practitioner: Implementing cooperative learning: A field study evaluating issues for school-based consultants. Journal of School Psychology, 32, 67-84.

Brown, I.S., & Felton, R.H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. Reading and Writing: An Interdisciplinary Journal, 2, 223-241.

Bruer, J.T. (1997). Education and the brain: A bridge too far. Educational Researcher, 26(8), 4-16.

Caine, R.N., & Caine, G. (1991). Making connections: Teaching and the human brain. Alexandria, VA: Association for Supervision and Curriculum Development.

Campbell, D., & Stanley, J. (1966). Experimental and quasi-experimental designs for research. Skokie, IL: Rand McNally.

Caplan, S.W., & Ruble, R. (1964). A study of culturally imposed factors on school achievement in a metropolitan area. Journal of Educational Research, 58, 16-21.

Carbo, M. (1988). Debunking the great phonics myth. Phi Delta Kappan, 70, 226240.

Chall, J. (1967/1983). Learning to read: The great debate. New York: McGraw-Hill.

Carnine, D. (1993, December 8). Facts over fads. Education Week, p. 40.

Carnine, D. (1995, May 3). Is innovation always good? Education Week, p. 40.

Clay, M.M. (1988). Studying developmental change with a successful intervention. Unpublished paper. (ERIC Document Reproduction Service ED 299 556).

Cohen, P. (1995). Understanding the brain: Educators seek to apply brain based research. Education Update, 37 [Online Document]. Available at http://www.ascd.org/pubs/eu/septu95.html .

Cook, T.D, & Campbell, D.T. (1979). Quasi-experimentation, design & analysis issues for field settings. Boston: Houghton Mifflin.

Cremin, L.A. (1964). The transformation of the school: Progressivism in American education 1876-1957. New York: Vintage Books.

Cuban, L. (1993). How teachers taught: Constancy and change in American classrooms, 1890-1990. New York: Teachers College Press.

Dare, M. (1997). Backgrounder #1: Learning to read. Organization for Quality Education. [Online Document]. Available at http://www.interlog.com.

Darling-Hammond, L., Griffin, G., Wise, A. (1992). Excellence in teacher education: Helping teachers develop learner-centered schools. Washington, DC: National Education Association.

Davis, R.A. (1943). Applicability of applications of psychology with particular reference to schoolroom learning. Journal of Educational Research, 37, 19-30.

Davis, W.A. (1948). Social class influences upon learning. Cambridge: MA: Harvard University Press.

Dewey, J. (1916/1963). Democracy and education. New York: Macmillan Publishing USA.

Druckman, D., & Bjork, R.A. (Eds.). (1994). Learning, remembering, believing: Enhancing team and individual performance. Washington, DC: National Academy Press.

Eason, S. (1992, January). Power assessment and the Delphi process. Unpublished paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX.

Edelman, G. (1987). Neural Darwinism: The theory of neuronal group selection. New York: Basic Books.

Ellis, A.K., & Fouts, J.T. (1993). Research on educational innovations. Princeton Junction, NJ: Eye on Education, Inc.

Ellson, D. (1986, October). Improving teaching productivity. Phi Delta Kappan, 111-124.

Engelmann, S. (1992). War against the schools' academic child abuse. Portland, OR: Halcyon House.

Engelmann, S., Becker. W.C. Carnine, D., & Gersten, R. (1988). The Direct Instruction Follow Through model: Design and outcomes. Education and the Treatment of Children, 11, 303-317.

Entwistle, N. (1981). Styles of learning and teaching. New York: Wiley.

Epstein, H. (1978). Growth spurts during brain development: Implications for educational policy and practice. In J. Chall and A. Mirsky (Eds.), Education and the brain: The 77th yearbook of the national society for the study of education, part II. Chicago: University of Chicago Press.

Foorman, B.R. (1994). Exploring connections among reading, spelling, and phonemic segmentation during first grade. Reading and Writing: An Interdisciplinary Journal, 6, 65-91.

Giacomia, R.M., & Hedges, L.V. (1982). Identifying features of effective open education. Review of Educational Research, 52, 579-602.

Glick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Glick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Good, T.L., & Brophy, J.E. (1986). Educational psychology (3rd ed). New York: Longman.

Groff, P. (1994, January). Rethinking whole language. Executive Educator, 16(1), 33-35.

Grossen, B. (1997). 30 years of research: What we now know about how children learn to read. A synthesis of research on reading from the National Institute of Child Health and Human Development commissioned by the Center for the Future of Teaching and Learning with funding support from the Pacific Bell Foundation. http://www.cftl.org/30years/30years.html

Hambleton, R.K., & Murphy, E. (1992). A psychometric perspective on authentic measurement. Applied Measurement in Education, 5, 1-16.

Hare, A.P. (1962). Handbook of small group research. New York: Free Press.

Healy, J.M. (1990). Endangered minds: Why children don't think and what we can do about it. New York: Simon & Schuster.

Heffernan, H. (1958). Evaluation--more than testing. National Education Association Journal, 47, 227-229.

Hilgard, E.R. (1939). The relation of schools of psychology to educational practices. California Journal of Elementary Education, 8, 17-26.

Hirsch, E.D. (1996). The schools we need and why we don't have them. New York: Doubleday.

House, E., Glass, G., McLean, L., & Walker, D. (1978). No simple answer: Critique of the Follow Through evaluation. Harvard Educational Review, 48, 128-160.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.

Johnson, J., & Immerwahr, J. (1994). First things first: What Americans expect from the public schools (a report from Public Agenda). New York: Public Agenda.

Johnson, J.E., & Johnson, K.M. (1992). Clarifying the developmental perspective in response to Carta, Schwartz, Atwater, and McConnell. Topics in Early Childhood Special Education, 12(4), 439-457.

Krathwohl, D.R. (1993). Methods of educational and social science research: An integrated approach. New York: Longman.

Lieberman, M. (1993). Public education, an autopsy. Cambridge, MA: Harvard University Press.

Lipsey, M.W., & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta- analysis. American Psychologist, 48, 1181-1209.

Lorge, I., Fox, D., Davitz, J., & Brenner, M.A. (1958). A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. Psychological Bulletin, 53, 337-372.

Madaus, G.F. (1994). Testing's place in society: An essay review of "Testing testing: Social consequences of the examined life" by F. Allen Hanson. American Journal of Education, 102, 222-234.

NAEYC [National Association for the Education of Young Children]. (1991 March). Guidelines for appropriate curriculum content and assessment in programs serving children ages 3 through 8: A position statement of the National Association for the Education of Young Children and the National Association for the Education of Early Childhood Specialists in State Departments of Education. Young Children, 21-38.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U.S. Department of Education.

Olson, L. (1997, January 22). Keeping tabs on quality. In Quality counts, an Education Week/Pew Charitable Trusts special report on the condition of education in the 50 states. Education Week, 16 (Suppl.), 7-11, 14-17.

Pascual-Castroviejo, I. (1996). Neuronal plasticity. Review of Neurology, 24, 1361-1366.

Paulu, N. (1988). What we know about phonics: Research in brief. Office for Educational Research and

Improvement. (ERIC Document 297-298).

Pirolli, P.L., & Anderson, J.R. (1985). The role of practice in fact retrieval. Journal of Experimental Psychology: Learning, Memory, & Cognition, 11, 136-153.

Pollard, R.R., Pollard, C.J., & Meers, G. (1994). Determining effective transition strategies for adjudicated youth with disabilities: A national Delphi study. Journal of Correctional Education, 45, 190- 196.

Proper, E.C., & St. Pierre, R.G. (1980). A search for potential new Follow Through approaches: Executive summary. Cambridge, MA: Abt Associates.

Regnier, P. (1996). Research, reason, truth, and education: Policy decisions and the intellectual life of schools. Unpublished manuscript.

Rosenshine, B. (1978). Review of teaching styles and pupil progress. American Educational Research Journal, 15, 163-169.

Sackman, H. (1974, April). Delphi assessment: Expert opinion, forecasting, and group process. (R-1283-PR). Santa Monica, CA: RAND Corporation.

Scheirer, M.A., & Kraut, R.E. (1979). Increasing educational achievement via self-concept change. Review of Educational Research, 49, 131-150.

Sears, S., & Keogh, B. (1993). Predicting reading performance using the Slingerland procedures. Annals of Dyslexia, 43, 78-89.

Shiffrin, R.M., & Schneider, W. (1977). Control and automatic human information processing, I: Detection, search and attention. Psychological Review, 84, 1-66.

Slavin, R.E. (1995). Cooperative learning: Theory, research, and practice (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Snow, R.E. (1992). Aptitude theory: Yesterday, today, and tomorrow. Educational Psychologist, 27(1), 5-32.

Snow, R.E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. Annual Review of Psychology, 43, 583-626.

Spaulding, F.E. (1903). The teacher's practical application of the results of child study. Journal of Pedagogy, 16, 34-42.

Stahl, S.A., & Miller, P.D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. Review of Educational Research, 59, 87-116.

Stanovich, K.E. (1996). How to think straight about psychology (4th ed.). New York: HarperCollins.

Steinberg, L. (1996). Beyond the classroom: Why school reform has failed and what parents need to do. New York: Simon & Schuster.

Stern, D. (1996). Active learning in students and teachers. Paris: Organization for Economic Cooperation and Development.

Stevens, R.J., & Slavin, R.E. (1995). The cooperative elementary school: Effects on students' achievement, attitudes, and social reactions. American Educational Research Journal, 32, 321-351.

Stone, J.E. (1996). Developmentalism: An obscure but pervasive restriction on educational improvement. Education Policy Analysis Archives, 4(8). Available http://olam.ed.asu.edu/epaa/v4n8.html

Strauss, H.J., & Ziegler, L.H. (1975). The Delphi method and its uses in social science research. Journal of Creative Behavior, 9, 253-259.

Sylwester, R. (1994). What the biology of the brain tells about learning. Educational Leadership, 51, 46-51.

Turner, R.L. (1989). The "great" debate: Can both Carbo and Chall be right? Phi Delta Kappan, 71, 276283.
U.S. Department of Education. (1996). NAEP Facts (National Assessment of Educational Progress). Washington, DC: Author.

U.S. Department of Health and Human Services. (1985). The impact of Head Start on children, families, and communities: Final report of the Head Start evaluation, synthesis, and utilization project, executive summary.

Washington DC: Author.

Vellutino, F.R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code- oriented versus whole-language approaches to reading instruction. Journal of Education Psychology, 83(4), 437443.

Walberg, H.J. (1990, February). Productive teaching and instruction: Assessing the knowledge base. Phi Delta Kappan, 470-478.

Walberg, H.J. (1992). The knowledge base for educational productivity. International Journal of Educational Reform, 1, 1-10.

Walker, M.H. (1996). What research really says. Principal, 75(4), 41, 43.

Watkins, C.L. (1988, July). Project follow through: A story of the identification and neglect of effective instruction. Youth Policy, 7- 11.

Wiersma, W. (1995). Research methods in education (6th ed.). Needham Heights, MA: Allyn & Bacon.

Willson, V.L. (1991). Performance assessment, psychometric theory and cognitive learning theory: Ships crossing in the night. Contemporary Education, 62, 250-254.

Winn, W. (1994, October). Why I don't want to be an expert sitar player. Educational Technology, 11-14.

Wolk, R.A. (1997, January 22). Quality counts, an Education Week/Pew Charitable Trusts special report on the condition of education in the 50 states. Education Week, 16 (Suppl.).

Woolfolk, A. (1998). Educational Psychology (7th ed.). Needham Heights, MA: Allyn & Bacon.

J. E. Stone is an educational psychologist and professor in the College of Education at East Tennessee State University. He also heads the Education Consumers ClearingHouse (http://www.education-consumers.com)

Andrea Clements is an associate professor in the Department of Human Development & Learning, College of Education at East Tennessee State University

The Education Consumers Consultants Network is an alliance of experienced and credentialed educators dedicated to serving the needs of parents, policymakers, and taxpayers for independent and consumer-friendly consulting. For more information, contact J. E. Stone, Ed.D., at (423) 282-6832, or write: professor@education-consumers.com