EFFECTIVE School Practices

VOLUME 14, NUMBER 1 PRICE $5.00 WINTER 1995

HANDBOOK FOR GRASSROOTS REFORM

Coming Issues:
"Reasoning and Creativity"
and
"What was that Project Follow Through?"

ADI

Philosophy of *Effective School Practices*

1. Teachers are responsible for student learning.
2. The curriculum is a critical variable for instructional effectiveness.
3. Effective teaching practices are identified by instructional research that compares the results of a new practice with the results of a viable alternative.
4. Experiments should not be conducted using an entire generation of Americans. The initial experimentation with a new practice should be small in scale and carefully controlled so that negative outcomes are minimized.
5. A powerful technology for teaching exists that is not being utilized in most American schools.

Effective School Practices is published quarterly by the Association for Direct Instruction. The mission of the Association for Direct Instruction, as stated in the by-laws, is to promote the improvement of educational methods.

The name *Direct Instruction* originated with the highly effective instructional model first developed by Zig Engelmann in Project Follow Through during President Johnson's Great Society legislation. Although the evaluation of Project Follow Through showed the Direct Instruction model to be far more effective than the other models on every identified outcome, education in America remained generally unchanged.

A few educators, impressed by the extraordinary results of the original Direct Instruction model and the programs that were developed as DI evolved, formed the Association for Direct Instruction in 1981.

Today, this organization is a vanguard in promoting school practices that have been validated as effective through the use of the scientific method in educational research.

The Association for Direct Instruction was incorporated in 1981 in the state of Oregon for educational purposes. ADI is a nonprofit, tax-exempt corporation under Section 501(c)3 of the Internal Revenue Code and is a publicly supported organization as defined in Sections 170(b)(1)(A)(ii) and 509(a)(1). Donations are tax-deductible.

A copy or summary of the current financial statement, or annual report, and registration filed by ADI may be obtained by contacting: ADI, PO 10252, Eugene, OR 97440 (503-485-1293). ADI is registered with the state of Oregon, Department of Justice, #79-16751.

ANNUAL SUBSCRIPTION RATES: $20 U. S.; $25 (U. S. currency or equivalent) Canada; $30 Europe; $40 airmail to Europe.
(ISSN 1068-7378).

e-mail: BGrossen@oregon.uoregon.edu

From the Field: Letters

Dear Reader,

An earlier issue (OBE and World Class Standards, Vol. 13, No. 3) was generally critical of top-down reform for its one-size-fits-all mentality. In this issue we feature reference materials and guides that we hope will help schools establish reliable bottom-up reform, particularly in the selection of the teaching methods and tools (textbooks, technology, media, software, and so on).

The first article by Russell Worrall and Doug Carnine describes in some depth the problem that we hope the rest of the issue will help solve: the irrationality of educational decision-making. Here the irrationality of education is contrasted with the rational decision-making processes in medicine and engineering. After reading this article, your confidence in decisions made by the educational bureaucracy should be thoroughly shaken.

Individual school communities that wish to use a more rational process will need to build that process themselves at the local level. The Handbook for Site Councils provides a road map for doing so. Your school may currently have a site council, or school improvement team of some sort, that either makes decisions of little import or makes important decisions without adequate information. This Handbook is a guide for changing that.

The brief overview of the Mattawan, Michigan school system decision-making process provides a working model for monitoring the implementation of the site council's decisions in a cooperative relationship between the school personnel and the school board / community.

The key to school improvement is selecting interventions that are likely to improve learning. This selection process requires that site councils obtain reliable information about what works, that is, site councils should select proven practices and tools. Reliable information is usually available in the form of research studies. Because research is often misused and abused, we have included a short guide for reading the kind of research that is most relevant for selecting teaching methods and tools. Such research reports the effects of the use of new tools on student learning.

As an example of the type of research that evaluates the effects of using various teaching procedures and tools and is, therefore, highly relevant to school decision-makers, we include a research report by Sara Tarver and Jane Jung. Readers may practice their research reading skills in reading this research report that compares the effects of two mathematics curricula on student learning.

Bonnie Grossen, Editor

To the editor:

Two years ago, the school district in my community implemented a whole-language reading program. Teachers were forced to use only a literature-based approach which had almost no phonics. There was no grouping of children to account for difference in needs. We are a district where disadvantaged children represent the majority. The program was too difficult for most of these children. Results were quite disastrous. Because children were having difficulty learning to read, the district attempted to "solve" its problem by classifying these children as handicapped. The present prevalence rate is very high (17%). Even before implementation of the new reading program, however, disadvantaged children were inappropriately classified. In addition, at one school, 62% of exiting kindergarten children did not meet criterion on a reading readiness test. The same reading readiness test was given to exiting grade 1 children who had received a year of reading instruction. Although the criterion was raised somewhat, 66% did not meet criterion.

There is some interest and great need to return to a more phonics-based approach. I am very interested in receiving information regarding your program.

From New York

Dear New York Reader:

Your district seems to be using irrational decision-making methods, as many school districts are. This handbook may be helpful for you. Your problems started with the selection of an unproven tool for teaching reading–whole language. As you search for a new tool, I would suggest looking for evidence that the new tool works well before you adopt it. Two schools might interest you–Wesley Elementary School in Houston, Texas, and Kreole Elementary in Moss Point, Mississippi. Both schools serve minority populations, but both schools achieve remarkable results in reading (Kreole recently made headline news for scoring second highest in the state). I would suggest visiting one of these schools, if you have the opportunity.

Ed.

ADI Materials Price List

Theory of Instruction (1991)
by Siegfried Engelmann & Douglas Carnine
Membership Price: $32.00 List Price: $40.00 ________

The Surefire Way to Better Spelling (1993)
by Robert Dixon
Membership Price: $8.75 List Price: $12.00 ________

Teach Your Child to Read in 100 Easy Lessons (1983)
by Siegfried Engelmann, Phyllis Haddox, & Elaine Bruner
Membership Price: $14 .95 List Price: $17.95 ________

Teacher Monitoring Program (1992)
by Colin Bird, Elizabeth Fitzgerald, & Margaret Fitzgerald
Membership Price: $15.00 List Price: $15.00 ________

Structuring Classrooms for Academic Success (1983)
by Stan Paine, J. Radicchi, L. Rosellini, L. Deutchman, & C. Darch
Membership Price: $11.00 List Price: $14.00 ________

War Against the Schools' Academic Child Abuse (1992)
by Siegfried Engelmann
Membership Price: $14.95 List Price: $17.95 ________

Subtotal ________

Postage & Handling:

If your order is:	P & H is:
$0.00 to $20.99	$4.00
$21.00 to $40.99	$5.50
$41.00 to $60.99	$7.00
$61.00 to $80.99	$8.50
$81.00 or more	$10.00

P & H ________

ADI Membership Fee ________

Total ________

Check enclosed. ☐
Bill my credit card. ☐ Circle one: Visa Mastercard

Number: ☐☐☐☐ ☐☐☐☐ ☐☐☐☐ ☐☐☐☐

Expiration date: ________

Signature: ____________________________________

Name (please print): ____________________________________

Address: ____________________________________

You may also phone in your order with Visa or Mastercard. **1-800-995-2464**.

Send to:
ADI
PO Box 10252
Eugene, OR 97440

Lack of Professional Support and Control Undermines Education: A Contrasting Perspective from Health and Engineering

Russell S. Worrall, O.D.
University of California, Berkeley
Doug Carnine, Ph.D.
University of Oregon

Confusion and conflict over academic and social goals for public education make it difficult for educators to move in concert toward any particular goal. In this frustrating context, teachers are expected to take responsibility for a rapidly expanding range of student performance levels to meet more demanding world-class standards. The growing diversity in America's student population further increases the complexity of the challenges educators face: 1 out of 10 children has a disability; over 1 out of 10 live in poverty; 1 out of 7 children over the age of 5 grows up speaking a language other than English.

However, a variety of dramatic educational successes with America's neediest students—Jamie Escalante in California, Marva Collins in Illinois, Thaddeus Lott in Texas—affirms the potency of education as a profession (Watkins, 1988). Moreover, America's advantaged students continue to perform at relatively high levels. A 1988 study of mathematics achievement of 13-year-olds (OECD, 1993) reported the highest performers to be from Taiwan, Iowa, North Dakota, Korea, Minnesota, the Soviet Union, Switzerland, Maine, New Hampshire, and Hungary. In a 1994 international mathematics competition for high-school students, the American students not only won but also received a perfect score, the only time this has happened in the 35-year history of the competition.

While these achievements of American students are encouraging, many teachers face students who experience serious learning problems. These teachers are expected to develop or to "fix" their own educational tools, such as textbooks and activity guides, so they will better meet the needs of students with learning problems. More recently, with the advent of decentralization, teachers are being asked to completely redesign their schools and participate in administering new programs, usually in their "spare" time. This expectation is unreasonable given professional training that teachers themselves rate as relatively ineffectual in comparison to the ratings other professionals give their own training.

Teachers are being asked to completely redesign their schools and participate in administering new programs, usually in their "spare" time. Teachers are led to believe that responding to these demands demonstrates their creativity and dedication.

Teachers are led to believe that responding to these demands demonstrates their creativity and dedication. Teachers are expected to be rugged individuals, able to surmount all obstacles. Why is it that no other profession places such a burden on its members? In health and engineering, professionals take on demanding responsibilities, but within a limited focus and only with extensive support and control from inside the profession, from the government, and from outside groups. In other professions, operating without such support and control would not be seen as creativity and dedication but malpractice.

The purpose of this essay is to contrast the professional support and control for teachers with that for other professionals and to illustrate the near impossibility of widespread educational reform without genuine professional support. Professional support includes:

- construction and acceptance of a common, accepted body of scientific knowledge about procedures and approaches that work, with recommendations for how those procedures and approaches can be implemented.
- pre-service education with a clear mission to teach prospective teachers when and how to select and implement validated procedures and approaches.
- educational materials, which if used as directed, have been demonstrated to be successful in teaching the desired educational objectives.
- valid and reliable measurement tools that inform teachers as well as parents and the community about the achievement levels of students.
- continuing on-the-job professional development, including the ongoing services of a consultant regarding validated educational interventions with recommendations of what and how they can be used to improve results.

Based on the present analysis, a primary target for reform must be on educational leaders who have failed to build a viable support system for teachers.

Only after the problems caused by insufficient professional support are understood and acknowledged can a fruitful discussion of remedies begin. These remedies entail building a professional support and control system, one that teachers need and deserve so they can better serve their students, particularly those who are in the greatest need of an effective education. Based on the present analysis, a primary target for reform must be on educational leaders who have failed to build a viable support system for teachers.

Rational Methods and the Art of Professional Practice

No profession will ever be totally scientific; there will always be craft and artistic components—"science in the service of the art of medicine." However, the key characteristic which has led to the problem-solving success of professions such as engineering and medicine is a reliance on *rational methods*. Rational methods include a knowledge base derived through the application of the scientific method and through the compilation of seemingly effective craft practices for problems that have not been scientifically investigated. The rational for the adoption of various practices and approaches is open to public scrutiny and incorporates available scientific findings and is logically organized. Over time, a knowledge base growing from an accumulation of research data and scientific evidence replaces dogma, faith, and faddism.

The industrial revolution and modern healthcare practices began with the questioning of the traditional teachings of classic Greek, Roman, and Arabic scholars. Rational methods, the scientific method in particular, for determining what works were implemented in numerous areas. In medicine, controlled clinical trials were conducted in which new methods were compared to existing techniques, a placebo, and simply doing nothing. With systematic observation and experimentation, new laws of motion and mechanics were developed providing engineers with the tools necessary to create products which have allowed man's productivity to exceed subsistence levels.

Trust in a profession is built on the assumption that a doctor is using methods *proven* safe and effective to varying degrees and that an engineer has chosen components for an aircraft that have been *demonstrated* to be reliable in relation to cost. A reliance on rational methods produces tools that enhance the probability of success for a specific problem. Replication of studies and generalization of models make the resulting tools more reliable and useful. A body of knowledge based on rational methods is dynamic, changing in response to new scientific data, creating tools that must hold up to the critical scrutiny of research. The resulting body of knowledge largely defines the internal, governmental, and external support and control that shape a profession and, to a large extent, set the benchmark for quality assurance (and thus status) for professions.

Internal Support and Control

The internal forces that have molded modern healthcare and engineering are based on a reliance on scientific methods and an individual depth and breadth of knowledge sufficient to manage complex problems. These are nurtured by university-based training and research programs and professional organizations.

Because of the relative unimportance of rational methods in education, professional support and control are based primarily on consensus among educational leaders.

In contrast, in education these controls and supports are underdeveloped, absent, or dogma driven, suggesting a preprofessional status. Because of the relative unimportance of rational methods in education, professional support and control are based primarily on consensus among educational leaders. They believe and feel strongly that certain practices or approaches should be used by teachers. The basis for consensus is belief and strong feelings. Beliefs and strong feelings also play a role in consensus-forming in other professions, but they dominate in education.

While consensus may be appropriate for setting standards that deal with values, such as what should be taught, what constitutes effective instructional standards is a matter of research and demonstration, not consensus among educational leaders. Without the constraint of rational methods, consensus in education more easily embraces educational fads, dogma, and ineffective or poorly implemented approaches.

While consensus may be appropriate for setting standards that deal with values, such as what should be taught, what constitutes effective instructional standards is a matter of research and demonstration, not consensus among educational leaders.

It is important to note that several prominent educators argue that education should not be viewed as a profession (Rowan, 1994, p. 4):

> Some analysts take the position that teaching is a form of craft work, that is, a set of well-established work practices grounded in the wisdom of accumulated practice rather than a highly codified and advanced scientific knowledge base (Huberman, 1993; Pratte & Rury, 1991). Others argue that in the absence of a well-developed science of teaching, teachers' work is best seen as a type of artistic endeavor that requires intuition and inspiration for successful performance (Eisner, 1978).

However, the availability of objective data does not necessarily preclude the application of "craft" knowledge or artistic endeavor in education. Teachers do not need to fear that they will become automatons simply because rational information is available to inform the*ir* teaching decisions.

Leaders of the educational research community express doubt that research should or can meet the needs of teachers and policy makers concerned with school improvement. A 1993 article in *Educational Researcher* summarized an electronic discussion forum on education priorities for the U.S. Office of Educational Research and Improvement (OERI) that ran over three weeks and produced dozens of contributions from the 700 or so participants in the forum. The opening remarks are revealing:

> Gene Glass: Educational research is a debate, an argument. . . . Some people expect educational research to be like a group of engineers working on the fastest, cheapest, and safest way of traveling to Chicago, when in fact it is a bunch of people arguing about whether to go to Chicago or St. Louis (p. 17).

The closing comments suggest that findings from educational research cannot be of much import (Glass, 1993):

> David Berliner: . . . What we do in school must necessarily constitute a weak treatment (p. 21).

Other researchers believe that research for developing and evaluating instructional materials is unprofessional. Shannon (1987), an internationally renowned expert in reading, put it this way, "The technical control of reading programs (the commercial reading materials) deskills teachers by supplying the goals, means, and evaluation of their reading instruction," (p. 321). This is analogous to claiming that engineers who use proven analysis and construction methods to design a truss for a bridge or doctors who use a clinically tested diagnostic and treatment regime are deskilled. In fact, most professionals rely on methods and tools that have well-established goals, means, and evaluation methods. Not using those methods and tools when appropriate constitutes malpractice.

Provisions do not exist to determine the safety and efficiency of the educational tools that are at the heart of our educational system.

Inadequate application of solid research on effective instructional tools and practices hamstrings educational reform. Provisions do not exist to determine the safety and efficiency of the educational tools that are at the heart of our educational system, in spite of the fact that, according to the Education Product Information Exchange, such tools are used from 75 to 90% of the 30 billion hours in which America's 40 million students are in school (Komoski, 1992). As Tyson-Bernstein pointed out with respect to textbooks:

> At present, there is no intellectually defensible method of verifying the effects of a book on learners because school officials have generally surrendered the task to the publishers, who use field trials as a sales strategy. Until school systems have a valid, economical method for testing books on students, their efforts to put student interests ahead of all others will be hampered (1988, p. 76).

Until school systems have a valid, economical method for testing books on students, their efforts to put student interests ahead of all others will be hampered.

The Social Studies Review (Sewall, 1992) contained an article that described how educational tools are developed and marketed. The article was prompted by the inaccuracies found in several U.S. history textbooks:

> The actual writers of textbooks are almost never the stated authors on the spine of the book. Such "names" are usually consultants, and the actual work is done by anonymous writers, too often of modest talent....
>
> Murray Giles, the editorial director for social studies and health at Glencoe, the high school imprint of Macmillan/McGraw-Hill, blamed time pressures . . . "we have to conduct market research, find an author, get the manuscript written and reviewed by experts, do the art work and photo-editing, and finally print the book . . . "(p. 12).

The chronology of the development process, described above by Murray Giles, does not even mention trying out educational tools with students to determine their efficacy.

Teachers are aware of the shortcomings of educational tools and methods. In the National Teacher Survey (1990) conducted by the Carnegie Foundation for the Advancement of Teaching, more than two-thirds of the nation's teachers listed improved instructional materials and supplies as essential for improving the quality of education in the United States. A study of 65 mathematics teachers in five states (Kelly, Dimino, Kameenui, & Carnine, 1993) found that two-thirds of the teachers felt their tools were inadequate for diverse learners, e.g., children of poverty, individuals with disabilities, and students with limited-English proficiency. Teachers were frustrated by the lack of time to make their own materials or even to modify existing ones. In the study, while more than 70% attempted to modify their books and materials, 80% said they did not have enough time to make the needed improvements.

Two-thirds of the nation's teachers listed improved instructional materials and supplies as essential for improving the quality of education in the United States.

The development of professional tools and methods is not the complete answer to successful educational reform. It is, however, one indispensable ingredient. With professional tools, it will become clearer what effects educators can and cannot produce.

Professional Organizations

Within most professions, there are organizations that foster quality assurance. General membership groups such as the American Medical Association, American Optometric Association, and The American Society of Mechanical Engineers promote professional development through educational meetings and publication of refereed journals. Board certification is also offered in various sub-specialties requiring a demonstration of knowledge and professional skill. Panels that review complaints of professional malpractice are a common component of these professional organizations. Peer review results in a consensus on the standard of care or professional expertise expected of the average individual in practice. Through advocacy, these organizations also promote legislation to clearly define the profession.

In education, there are organizations for professors that have a research focus, e.g., American Educational Research Association. However, the major general membership organizations for teachers, such as the American Federation of Teachers and the National Education Association, have functioned primarily as bargaining units with an emphasis on advocacy for teachers' rights and compensation, not professional development. However, they are paying increasing attention to professional issues, through AFT's Educational Research and Dissemination's group and NEA's Professional Standards and Practice.

Several national curriculum organizations exist for the basic curriculum areas such as math, science, English, and reading. They develop academic goals and standards for what should be taught and learned in our public schools. The National Council of

Teachers of Mathematics (NCTM), the International Reading Association (IRA), and the National Council of Teachers of English (NCTE) have worked diligently to establish consensus around content standards.

There is an interesting twist to their activities, however. They have also busied themselves with specifying instructional standards; i.e., the methods, approaches, and assessment procedures that should be used. While consensus based on beliefs and values is reasonable for identifying what content should be taught, beliefs and values are not sufficient for setting standards about how to teach.

The intent of these national organizations, such as the National Council of Teachers of Mathematics (1989), is to "... ensure that the public is protected from shoddy products" (NCTM, 1989, p. 2). The Standards drew an analogy between the evidence of the instructional effectiveness of mathematics programs with the kinds of evidence used by the Food and Drug Administration to establish minimum quality criteria for the distribution of drugs. The NCTM Standards (1989, p. 2) went on to assert that: "It seems reasonable that anyone developing products for use in mathematics classrooms should document how the materials are related to current conceptions of what content is important to teach and should present evidence about their effectiveness."

Bishop (1990), an ardent supporter of the Standards, pointed out that authoritative opinion was apparently a more compelling force than empirical research in defining effectiveness. He argued that while the authoritative opinion of mathematics educators is the principal basis for establishing goals of mathematics education, authoritative opinion does not provide the same kind of support as empirical research for the development of instructional practices. Earlier, the NCTM's Research Advisory Committee (RAC, 1988) had raised the same concern:

> Although there is no reason to expect a solid research base for every suggestion made in the document, the draft version did not distinguish those recommendations that were well-grounded empirically or theoretically from those that were based more on the informed judgment or personal opinions of the authors or that were drawn from examples and experience available in other countries (p. 339).

In fact, the Standards document (1989) included this suggestion: "... the establishment of some pilot school mathematics program based on these standards to demonstrate that all students—including women and underserved minorities—can reach a satisfactory level of mathematics achievement" (NCTM, 1989, p. 253). It is difficult to imagine the Food and Drug Administration, chosen by the NCTM as a model for its Standards, approving a drug and afterwards proposing "the establishment of some pilot" to see if the drug helped or harmed people.

States also devise their own standards. California implemented its new "whole language" standards in 1988 to provide a quality education for all its students, including diverse learners. It then conducted school audits for accountability purposes. By the time of the fourth-grade reading assessment, almost 90% of the California fourth-grade teachers reported that they heavily emphasized the California whole language standards. Yet the California fourth-graders scored near the bottom of all participating states (Education Week, September 22, 1993). Stanford's Michael Kirst commented, "We almost beat Mississippi—but not quite. For California to say that is just devastating" (Oregonian, October 27, 1993, p. A01). Proponents of California's whole language standards point to a growing number of diverse learners to explain the state's poor performance, even though a primary purpose of setting the standards in the '80s was to meet the needs of diverse learners. (This argument does not account for the fact that white fourth-graders in California also scored near the bottom when compared to white fourth-graders across the U.S.).

California also used consensus among educational leaders to adopt instructional standards in mathematics in 1985. The next set of mathematics instructional standards for California (1990 draft) reported on the success of the 1985 standards: "We have a growing body of experience from teachers who have worked to achieve the goals of the 1985 Mathematics Framework, and have found that it is possible to avoid the pitfalls described above and to engage students in meaningful work." This item from the 1985 Standards illustrates meaningful work: "Write a set of directions for a younger student, explaining how to add 2/5 and 1/3. Then use a picture and write an explanation as to why you add fractions the way you do."

However, the claim of progress is contradicted by actual test results. In the spring of 1991, the National Assessment of Educational Progress results were released; only 1 in 12 California eighth-graders could even add two fractions, such as 2/5 and 1/3, let alone explain how to teach another student to add two such fractions. California scored in the lowest third of the participating states. How did the State Department of Education in California respond? Francie Alexander, who was acting Director of the Curriculum, Instruction and Assessment Division in 1985 and Associate Superintendent in 1991, was quoted in Newsweek (June 17), "We've all been led to believe we were above average." It would be interesting to know who misled Ms. Alexander.

Many reform proposals from curriculum organizations promote untested instructional procedures.

More recently, the National Center for the Learning and Teaching of Elementary Subjects (NCLTES) observed the extent to which 24 teachers incorporated the California mathematics standards into their mathematics teaching. A NCLTES author commended one teacher's attitude "of exploration and invention, conveying the idea that all students can learn, enjoy, and use mathematics" (Heaton, 1992, p. 155). Unfortunately, the teacher told her students to multiply length times width to find perimeter, and to multiply feet times yards when calculating volume. The second teacher was less knowledgeable about the instructional recommendations of the standards, but when she attempted to implement the suggestions, she mistaught averages to her students. The final two teachers, who seemed to have a better grasp of instructional practices and content for elementary school mathematics, produced the greatest student achievement, according to traditional measures. However, these teachers were criticized for not following the standards more closely. Yet the NCLTES authors conceded that they were not sure what teachers should do to meet the instructional standards. Furthermore, the NCLTES authors failed to record what, if anything, the students they observed had learned. In short, many reform proposals from curriculum organizations promote untested instructional procedures.

It is important to note that the public is not hostile to genuine reform where students clearly benefit; however, the public is tiring of rhetoric, questionable practices, and meager results, as suggested in the NCLTES study. In some cases, public hostility erupts. For example, controversy over the adoption of a nationally recognized innovation split the community of Littleton, Colorado. Littleton High School, in becoming a national model for transformational outcomes-based education, replaced academic course requirements with performance requirements, including several that were social and personal in nature, which drew the ire and disdain of many in the community. For example, a student who exhibits one of several relatively minor infractions on a performance assessment would not be able to graduate from high school. To prevent a student from graduating because of one of the following behaviors seems very unreasonable: . . . "takes sides in a conflict or is defensive." . . . "doesn't let the other person finish before expressing own ideas." . . . "uses name calling or makes accusations." . . . "must be guided or procrastinates." . . . "reacts emotionally to criticism, or is not willing to evaluate self constructively." . . . "resists authority." . . . "does not set realistic goals." . . . "makes snap decisions with little thought as to outcome."

The innovations in Littleton failed in part because of an overreliance on consensus among educational leaders in Littleton and at the national level, with insufficient attention to rational methods for testing the innovations.

More than half the juniors at Littleton High School did not believe they would be able to graduate. No Chicano student believed graduation was possible. Parents could not even find out how their children stood with respect to graduation requirements. Not surprisingly, this affluent community elected a slate of board members that returned to conventional high school course requirements for graduation. The innovations in Littleton failed in part because of an overreliance on consensus among educational leaders in Littleton and at the national level, with insufficient attention to rational methods for testing the innovations.

Educational leaders react in anger when their consensus is rejected. Rexford Brown (1994), a senior fellow at the Education Commission of the States (E.C.S.), which serves as the education think tank for the nation's governors, wrote a newspaper column on the Littleton's decision to abandon transformational outcome-based education:

> What we are witnessing in Littleton is the triumph of small-town fear over common sense and hope. . . . The politics of fear always become the politics of intolerance. And as the uncompromising politics of intolerance continue in the Littleton Public Schools, the district moves steadily toward becoming the Little Town Public Schools—little in imagination, little in daring, little in accomplishment.

This response from a professional agency, which treated genuine community concern about educational results with scorn, does little to engender support for educational innovations. Ron Brandt (1995), in the Association of Supervision and Curriculum Development's (ASCD) February issue of Educational *Leadership*, voiced the lack of confidence the public has in educational leaders: ". . . the majority of people want 'first things first' meaning safety, order, and 'the basics.' As for practices ASCD supports, such as whole language approaches to teaching reading and writing, heterogeneous grouping, and authentic assessment, people oppose or are skeptical about them" (p. 3). The public wants results but is being served innovation. Of course, innovation has its place; but it's not a complete meal.

Depth and Breadth of Knowledge

A professional training program introduces rational methods and tools and emphasizes the art of employing these methods, which requires depth as well as breadth of knowledge. For example, breadth involves knowing the limits of the profession and the capabilities of associated professions so as to make meaningful referrals. Awareness of ineffective methods is also important so that inappropriate methods or materials are not employed or recommended.

Course work in scientific and statistical methods is also important. A professional must be able to critically evaluate claims from the lay press and in the professional literature. Being able to differentiate between a paper based on anecdotal evidence or an author's theory from a well-conceived and executed study is a key skill. Also, professionals are taught to be very careful in drawing conclusions from evidence.

Being able to differentiate between a paper based on anecdotal evidence or an author's theory from a well-conceived and executed study is a key skill.

In education, however, conclusions are often drawn that misinterpret the information that is available. For example, a research synthesis on grouping by ability (Gamoran, 1992) stated in the first paragraph that typically high-performing students benefit from fixed ability grouping while low-performing students fall further behind. The second paragraph stated that flexible ability grouping according to skills may raise achievement for both high and low-performing elementary students. The third paragraph recommended dramatically curtailing the use of ability grouping! The author's conclusion ignores his own findings about the benefits of flexible grouping based on skills.

The use of a common terminology is central to professional communication. In the health professions, resources such as a medical dictionary or, in engineering, specialty reference works such as The Dictionary of Visual Science ensure that a professional from New York understands what her counterpart in California is writing about. Educational literature, on the other hand, is fraught with conflicting or vague terms that seem to change with each author. There is no standard such as a dictionary of educational science or reading science to provide for clear communication of theories or research results. The use of vague terms in education is illustrated by Chall (1991) in describing whole language:

> . . . views whole language in opposition to basal readers, not phonics. Others place whole language on the side of reading "whole books," not short selections in basals; still others stress that whole language means no teaching of skills, as opposed to doing so in basal reading programs. For others, it means empowering teachers to teach reading as they wish. For still others, it means integrating the teaching of reading with writing, speaking, and listening. And for a growing number, it means a philosophy of education and of life, not merely a method of teaching reading (Goodman, 1986).

Another example is outcome-based education. While Littleton, Colorado's version of transformational outcome-based education wrought trauma, a very different version of outcome-based education brought success to Arlington Heights, an affluent suburb outside of Chicago (Fitzpatrick, 1991). These two versions of outcome-based education actually contradict each other, reflected in a schism that has produced two separate educational "movements." These very different versions of outcome-based education illustrate the problem of vague, imprecise language.

Professionals must also recognize faddist, fringe, or fraudulent methods. Worrall (1990) cites examples of educators, either unwittingly or overtly, promoting worthless methods or products that have been promoted in the lay literature, such as colored lenses as a treatment for reading disability or subliminal tapes suggesting new behaviors to the subconscious. This concern over faddism is expressed by a prominent educator:

> I fear the bandwagon. I fear the proselytizers (p. 25). . . . Most schools move from innovation to innovation ("We are doing whole language, or cooperative learning, or curriculum integration") and define success as the implementation of the latest innovation. This is nonsense. What difference does any

innovation make if a school cannot determine effects on kids? (Glickman, 1992, p. 26).

Students, parents, and the public must be able to trust the reasoning behind a teacher's recommendations, but to do so, educators must rely on proven, effective tools. An extreme example of faddist practice occurred when the Del Norte Unified School District in Northern California elected to participate in a study of Applied Kinesiology and Neural Organization Technique (AK/NOT), as a means of improving the learning of special education students (Worrall, 1990). The theory states that learning disabilities are caused by misalignment of certain bones of the skull that can be corrected through AK/NOT manipulation. School district staff recruited parents to allow their children to participate in the AK/NOT treatment. A graphic example of the therapy is offered by the parent of a six-year-old patient with a speech delay:

> Two doctors worked on him very intently as I sat across his legs to help hold him down. They focused their attention on this head and mouth areas. They were applying such tremendous pressure to Donald's skull and the roof of his mouth that they would break into a sweat and their bodies would just shake with the force of their exertion. All during this so-called treatment, Donald was screaming and struggling to get away from these doctors.

In October of 1987, four parents of children receiving the AK/NOT treatment complained. The district did not respond. In March of 1988, the superintendent and three board members were defeated in a recall election. Several parents sued the district, eventually winning a settlement in the hundreds of thousands of dollars. Sadly, before the AK/NOT treatment began in Del Norte, The Utah Chiropractic Association had branded AK/NOT as hucksterism, and a Third Judicial District Court had issued a restraining order barring the use of AK/NOT in Utah. Possibly the most sobering finding is this:

> . . . unfortunately, the Del Norte School District and this study itself will be cited by promoters of the AK/NOT program throughout the world to support the marketing of this questionable approach to the treatment of learning disorders and neurological disease (Worrall, 1990, p. 48).

Implications for Internal Support and Control

The relatively weak role for rational methods undermines the potential contributions of professional organizations to improving teaching and learning. Increasing the acceptance and use of empirical methods in education would provide needed checks and balances to consensus, as is the case in other professions. Changes in values and use patterns concerning rational methods in education will come only through governmental controls and external controls, which are discussed in the next two sections.

Governmental Support and Control

State licensing boards enforce the body of legislation that defines the limits of a profession. For example, doctors and engineers have a defined scope of practice in which they operate. The boards also test individuals to assure they meet minimum standards of competence. Litigation around malpractices refines the "standard of care" for a profession. Product liability litigation also holds the tools that professionals use to rigid standards of safety and efficacy.

Increasing the acceptance and use of empirical methods in education would provide needed checks and balances to consensus, as is the case in other professions.

In education, malpractice, product liability, and governmental monitoring of tools are viewed from a very different perspective. Teachers cannot be held liable for professional malpractice, because the courts do not view teaching as a profession as there is not an agreed-upon body of knowledge related to "generally accepted professional practices."

Even when there exists legislation designed to exert some control over standards, educators are often able to by-pass or sabotage the intent of the law. As noted earlier, in 1988 the California State Board of Education mandated instructional standards that included a specific untested approach for teaching beginning reading. In making this mandate, the California State Board of Education explicitly refused to comply with a 1976 law requiring that educational tools be tried out with students and revised before adoption. For explicitly ignoring the 1976 law and other reasons, a judge ruled that the State Board's procedure allowing untested curricular methods to be adopted was illegal (Long, James L., Judge of the Superior Court, 1989). In December of 1991, in a 39-page ruling, the California Court of Appeals upheld every aspect of the initial ruling; subsequently, the California Supreme Court agreed with the Appeals Court.

In response, Senate Bill 1859 was passed on May 14, 1992, essentially repealing the 1976 law and serves a pattern in auto accidents that seems to be statistically related to one make of automobile or to a particular highway design, the DOT will fund research or require the manufacturer to investigate the problem. The FDA continues to monitor the effectiveness and safety of drugs after approval and will restrict use or remove from the market a drug whose performance is deficient.

The National Center for Educational Statistics and the Regional Educational Laboratories are far less potent.

Data Dissemination

Agencies such as the National Institutes of Health and the Food and Drug Administration routinely advise practitioners of important findings and possibly hazardous procedures and medications through regular publications. Letters or bulletins are used to alert doctors or engineers to a recently identified and potentially dangerous problem with a product or chemical (drug) that is currently on the market.

The National Library of Medicine and its Medline database is a potent tool for data dissemination. There are also several proprietary databases available to physicians and engineers which allow them to access a worldwide pool of information.

When a teacher requested data on *Mathematics Their Way,* a primary-grade program viewed as embodying the instructional standards of the National Council of Teachers of Mathematics, a publishing representative responded in this way: "Thousands of teachers are using both *Mathematics Their Way* and *Mathematics a Way of Thinking* in their classrooms. Teachers using these methods and materials are able to see growth in understanding take place, which cannot be measured by standardized tests and formal research."

If a teacher has a question about the efficacy of a textbook or educational tool, the process of collecting objective information is convoluted at best. Most publishers do not collect research data on the effectiveness of the tools they sell. Some publishers do not feel such research is necessary, or they confuse data obtained from research with personal testimony and opinion on effectiveness. When a teacher requested data on *Mathematics Their Way*, a primary-grade program viewed as embodying the instructional standards of the National Council of Teachers of Mathematics, a publishing representative responded in this way: "Thousands of teachers are using both *Mathematics Their Way* and *Mathematics a Way of Thinking* in their classrooms. Teachers using these methods and materials are able to see growth in understanding take place, which cannot be measured by standardized tests and formal research. . . . The most frequently made comment by teachers is that using *Math Their Way* and *Math a Way of Thinking* creates attitudinal changes in students. For these teachers mathematics becomes an exciting area of study—one that teachers and students enjoy."

Implications for Governmental Support and Control

Governmental agencies promote problem-solving and problem identification through licensing boards, the courts, and a variety of federal agencies. Licensing and agency actions also support and control education. In education, the problem is that licensing boards and agencies typically rely on consensus while minimizing the role of rational methods.

One course of action is for educational agencies to learn from agencies that work with other professions. An interagency government task force of other professionals and educators might suggest how education could utilize processes from other agencies for identifying and solving problems through research and regulations. For example, the testing and validating of tools could be compared in education and medicine, with the FDA serving as a starting point. Medline could serve as a model for making educational research information accessible and useful to teachers. Such a task force could be a starting point for infusing rational methods in education.

External Support and Control

All professions are influenced by external pressures beyond government. The primary external force driving the system is marketplace dynamics—the demand by patients, clients, students, parents, or companies for a quality service or product.

Third-Party Entities

Though individuals or companies purchasing services and products ultimately shape the quality of a service provided by a profession, many consumers rely on commercial or third-party enterprises such as a hospital or engineering firm that demand a specific level of competency of their professional staff. Hospitals, health maintenance organizations (HMOs), and insurance companies monitor standards of care. Audits of patient records and review committees monitor a doctor's patient care practices (such as surgical mortality as compared to the national average for a similar procedure). Independent accrediting organizations are often used by institutions to provide assurance that the quality of their program and uniformity of care or service are at a high level.

In education, school reviews are usually carried out by state agencies. In most states, these reviews focus on whether tools and practices endorsed by a consensus of educational leaders are being implemented. This problem is illustrated by comments of the Colfax High School (California) staff in a self-study report (Spring, 1994), made in preparation for an accreditation visit.

"With the willingness to try several new concepts and programs promoted by the school and the District, many staff members are feeling stress from so many changes with few links to the 'new programs' of the past." "Maintaining current programs and practices appears to and may actually conflict with the newer ideas." "The calls for reform have been so numerous that staff often feel the school is in need of focus and a unified approach." "There is a feeling among some staff members that the district improvement initiatives are unrealistic without addressing structural elements such as large class sizes and large counseling loads" (p. 29-32).

The Accreditation Committee, consisting of California educators and administrators, recommended: "That the district, school administration, and staff develop clearly defined roles and policies in a timely manner for effectively implementing the new reform models, Dr. William Glasser's Quality Schools, Reality Therapy, and the Theory of Outcome Based Education."

Neither the district nor the accreditation panel's final report questioned the efficacy of the many reforms and teaching tools being simultaneously introduced at Colfax High School. Much effort was made to use the current educationally "correct" language though the deeper understanding and long-term implications of these state-mandated changes remains controversial.

In some states, school reviews focus more on measures of student learning. In these states, education agencies tell about deficiencies but present few credible solutions. A "blame the victim" attitude prevails. In 1988, Alessi reviewed school psychologists' files on 5,000 students with learning problems. Out of these 5,000 cases, an astonishing 5,000 were deemed to be the fault of the student or the student's family. It is difficult to conceive that out of 5,000 learning problems, not one was due to a professional shortcoming. Medicine has an entire field of study, iatrogenics, which studies symptoms, maladies, and disorders that can be attributed to inadequate or incorrect practice of medicine.

Special Interest Groups

Special-interest advocacy groups such as the American Automobile Association (AAA) or the American Cancer Society are a potent force to influence quality in a specific professional sector. These organizations provide in-depth consumer education so that intelligent, reasonable choices can be made in the selection of professional services or products. They can also promote and fund independent research. A major role in quality assurance is the ability of a special-interest group to apply the aggregate weight of its members in advocacy before governmental and professional organizations to effect change. The AAA can influence automotive and highway design through its national activities. The American Cancer Association is active in funding cancer research and educating cancer patients about the relative effectiveness of various treatment options.

Special interest groups such as the American Cancer Society and independent organizations such as the Consumers' Union, Underwriters Laboratories, Better Business Bureaus, and the National Council Against Health Fraud either do not exist in education (except for parents of children with disabilities) or do not target the quality of student learning.

The largest special interest group in education is the Parent Teacher Association, which is active in education but does not address issues of quality. There are a number of special interest groups for parents of children with disabilities. These parents gain information, social support, and political power through membership in these groups. Even for students with disabilities, the issues often have to do with access to services and settings rather than the quality of teaching and learning.

Independent Organizations

Autonomous organizations such as Consumer's Union and Underwriters Laboratories (UL) exist to provide consumers with objective information and reporting on products and services. Through publications, such as *Consumer Reports* magazine, evaluations of new and existing products are disseminated in a format that is readily available to the lay consumer. They routinely publish objective reports on subjects from chiropractic care to contact lenses and from automobile tires to personal computers. Better Business Bureaus act as a registry for adverse complaints about products or services and as such are a resource for the consumer when selecting a service or product and a secondary check to governmental agencies on quality assurance. These consumer "watchdog" activities encourage professionals to maintain reasonable quality levels often above the minimum set by legal statutes.

Healthcare, as is education, is prone to faddism, misinformation, fraud, and quackery. The National Council Against Health Fraud (NCAHF) was formed in 1977 to combat this tendency to abandon reason in the marketplace. NCAHF's membership is composed of healthcare professionals, attorneys, and consumers all sharing a common concern for the quality of the healthcare system. The emphasis is on identifying and combating misinformation about healthcare practices and devices. The goal is to help consumers protect themselves from ineffective practices through direct education, and as a central, reliable resource for media and governmental agencies.

There are no independent organizations dealing with effectiveness in the educational arena.

Implications for External Support and Control

The starkest contrast between education and other professions exists in the area of external controls. Special interest groups such as the American Cancer Society and independent organizations such as the Consumers' Union, Underwriters Laboratories, Better Business Bureaus, and the National Council Against Health Fraud either do not exist in education (except for parents of children with disabilities) or do not target the quality of student learning. Hopefully, educational versions of Consumers' Union, Underwriters Laboratory, Better Business Bureau, etc., will grow in number and influence. Foundations could take the initiative in funding existing organizations such as Consumers' Union to conduct a feasibility study of an initiative in education. As was the case for governmental agencies, educators will need to work closely and for an extended period of time with professionals from these other organizations before meaningful reforms will occur.

Conclusion

Support and control within the education profession will change only in response to changes in governmental agencies and the creation of third party and independent organizations. Those changes will require close and extended collaboration between educators and other professionals. Moreover, the impetus for building these supports and controls must come from outside the education profession, as have all fundamental reforms in American education (Cuban, 1994).

Two hundred years ago, medicine was patient-centered. According to the dogma of the time, the doctor asked the patient what he or she thought would be a good remedy!

This analysis would be exceedingly discouraging were it not for the fact that all the more mature professions were once at the stage education is today. Two hundred years ago, medicine was patient-centered. According to the dogma of the time, the doctor asked the patient what he or she thought would be a good remedy! When a practitioner doesn't have effective tools for solving problems, the practitioner defers to the client. It's no wonder that Voltaire described medicine as the art of amusing the patient while nature cures the disease.

The scientific revolution will be no less challenging for education than it was for biology and medicine.

One of the revolutionaries who ushered in modern medicine, Ambroise Paré, did so by moving medicine from dogma to science. Several hundred years ago, the standard treatment for battle wounds was boiling oil. During one battle, Paré ran out of boiling oil. To the rest of his patients, he administered salve. This was not remarkable. What set him apart from his colleagues and their dogma was what he did next. He actually went to visit his patients to see if there were differential effects for boiling oil and salve (Haggard, 1929). The evaluation of different approaches helped medicine on the road to science. The road was rocky, however. As Dr. Haggard pointed out, physicians continued to use boiling oil:

> Paré's medical writings were in the vernacular; he was not a scholar; the medical writings of all other prominent medical men of the time were in Latin. The organized physicians of Paris found therein an excuse for attacking the works of Paré and attempting to prevent their publication: they cited not only his ignorance, "a man very impudent and without any learning," but also that in his teachings he departed from the established practices of the ancients (p. 39).

The scientific revolution will be no less challenging for education than it was for biology and medicine. Claude Bernard, considered to be one of the fathers of modern medicine, was accused of reducing biology and the mystery of life to the banality of a machine. He gained immortality in *The Brothers Karamazov* when a character shouted "Bernard" as a term of derision for the mechanistic spirit of science. But as Campbell (1986) pointed out:

> Bernard's unique contribution to science and thought was to show that life is not governed by a collection of laws that fit together according to human logic, but by laws that need to be looked at in the light of nature's logic, which is quite a different thing. What seems absurd to us may not seem absurd to nature (p. 46).

Life is not governed by a collection of laws that fit together according to human logic, but by laws that need to be looked at in the light of nature's logic, which is quite a different thing.

Education leaders have a responsibility to provide professional supports for teachers and students. Inadequate and insufficient professional support helps explain why so many attempts to improve student learning have failed. One way to understand the predicament teachers face in attempting to improve learning through reforms such as the restructuring of schools is to look back to medicine at the turn of the century. Physicians lacked the tools they have today—hygienic practices, antibiotics, X-rays, etc. To have demanded better results from them by reorganizing the hospitals in which they worked or by promoting competition among doctors through greater choice would have produced meager results and great frustration. Educational reform is in a similar situation today. There is no shortcut or substitute for the application of rational methods and the development of supports and controls from within the profession, from government, and from external groups.

References

Alessi, G. (1988). Diagnosis diagnosed: A systemic reaction. *Professional School Psychology, 3*(2), 145-151.

Bishop, A. (1990). Mathematical power to the people. *Harvard Educational Review, 60*(3), 357-369.

Brandt, R. (1995). Overview. *Educational Leadership, 52*(5), 3.

Brown, R. (1994). Small-town fears making Littleton a 'Little Town.' *Rocky Mountain News*, Mon., Feb. 28.

Campbell, J. (1986). *Winston Churchill's afternoon nap*. New York: Simon and Schuster.

Chall, J. S. (1991). American reading instruction: Science, art, and ideology. *All language and the creation of literacy*. Baltimore, MD: The Orton Dyslexia Society, pp. 20-26.

Carnegie Foundation for the Advancement of Teaching (1990). *The condition of teaching: A state-by-state analysis* (National Teacher Survey). Princeton, NJ: Author.

Cuban, L. (1994). Special education in the context of special education reform efforts. Paper presented at the 1994 OSEP Project Directors Meeting, Washington, DC.

Curriculum Development and Supplemental Materials Commission (1988, June). 1988 English-language arts adoption—reading. Sacramento, CA: California State Department of Education.

Feynman, R. P. (1985). *Surely you're joking, Mr. Feynman*. NY: W. W. Norton & Co.

Fitzpatrick, K. A. (1991, May). Restructuring to achieve outcomes of significance for all students. Educational Leadership, pp. 18-22.

Gamoran, A. (1992). Is ability grouping equitable? *Educational Leadership, 50*(2), 11-17.

Glass, G. (1993, Aug.-Sept.). Research news and comment—A conversation about educational research priorities: A message to Riley. *Educational Researcher, 22*(6), 17-21.

Glickman, C. (Sept., 1992). The essence of school renewal: The prose has begun. *Educational Leadership, 50*(1), 24-27.

Goodman, K. (December, 1986). *Examining our assumption about the relationship between teaching and learning*. Special discussion session presented at the annual meeting of the National Reading Conference, Austin, TX.

Haggard, H. W. (1929). *Devils, drugs, and doctors: the story of the science of healing from medicine-man to doctor*. New York: Harper & Brothers.

Heaton, R. M. (1992). Who is minding the mathematics content? A case study of a fifth-grade teacher. *Elementary School Journal (93)*2, 153-162.

Kelly, B., Dimino, J., Kameenui, E. J., & Carnine, D. (1993). *Teachers' assessments of the adequacy of basal mathematics textbooks for the needs of diverse learners*. (Technical Report). Eugene: University of Oregon, National Center to Improve the Tools of Educators.

Komoski, K. (1992, March 18). A testimony given by Kenneth Komoski, Director of the Education

Products Information Exchange, to the U.S. Congress Select Committee on Education in Washington, D.C.

Long, James L., Judge of the Superior Court. Statement of Decision, No. 361906, Dept. 14, in the Superior Court of the State of California, in and for the County of Sacramento, November 14, 1989.

National Center to Improve the Tools of Educators (1993). Survey of teachers' perceptions of new educational tools and teaching approaches. Eugene: University of Oregon.

National Council of Teachers of Mathematics, Commission on Standards for School Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

OECD (1993). *Education in the States and nations: Indicators comparing U.S. states with OECD countries in 1988*. Pittsburgh, PA: U.S. Government Printing Office.

Research Advisory Committee (1988). NCTM curriculum and evaluation standards for school mathematics: Responses from the research community. *Journal for Research in Mathematics Education, 19*(4), 338-344.

Report to the National Education Goals Panel (1993, November). Promises to keep: Creating high standards for American students. Goals 3 and 4 Technical Planning Group on the Review of Education Standards (94-01), p. 14.

Rowan, B. (1994). Comparing teachers' work with work in other occupations: Notes on the professional status of teaching. *Educational Researcher, 23*(6), 4-21.

Sewall, G. (1992). *Social Studies Review*, No. 9, Spring Issue.

Shannon, P. (1987). Commercial reading material, a technological ideology, and the de-skilling of teachers. *Elementary School Journal, 87*(3), 307-327.

Staff (1994, Spring). *Pursuing excellence: Joint WASC/CDE self study report.* Colfax High School, Colfax, California.

Tyson-Bernstein, H. (1988). *A conspiracy of good intentions: America's textbook fiasco.* Washington, DC: The Council for Basic Education.

Watkins, C. L. (1988). Project Follow Through: A story of the identification and neglect of effective instruction. *Youth Policy*, Education section, July, pp. 7-11.

Worrall, R. S. (1990). Detecting health fraud in the field of learning disabilities. *Journal of Learning Disabilities, 23*(4), 207-212.

Worrall, R. S. (1990, Fall). Neural organization technique—treatment or torture? *Skeptical Inquirer, 15*, 40-50.

NEW!

Nominate a Student for:

The Wayne Carnine Outstanding Student Improvement Award

$200
in cash will be awarded
to the *student* selected as the 1996 recipient

Plan Now!

Doug Carnine and a committee of ADI Board Members will select the 1996 winner in the Spring of 1996. Be prepared to describe the outstanding academic and / or behavioral improvement of the school-aged student you may nominate.

Also, contributions to the Wayne Carnine Outstanding Student Improvement Memorial Fund can be sent to:

Carnine Award
ADI
PO Box 10252
Eugene, OR 97440

A Handbook for Site Councils to Use to Improve Teaching and Learning

Doug Carnine
National Center to Improve the Tools of Educators
University of Oregon

Abstract: *Smart schools do not have to embrace any particular approach, philosophy, or innovation; they do however, have to embrace a rational, decision-making process. Many of the problems in education can be traced to the lack of a process to provide teachers with effective tools and practices and with adequate support in their use. This handbook describes a four-stage process for creating smart schools: (a) setting improvement goals reflected in measures acceptable to the school and community, (b) defining the scope of the improvement plan, (c) identifying approaches (i.e., tools and practices that are effective, sustainable, accountable, equitable, and cost efficient); (d) planning and managing the implementation of the selected approach.*

The assumption underlying the smart school process is that improvements in learning will result from improvements in what goes on in classrooms. Smart schools use information to select and implement effective tools and practices accompanied by the necessary organizational support and professional development. Without some form of a smart school process, innovations such as restructuring will not necessarily produce the desired improvements in student learning.

Stage 1. Setting Improvement Goals

The process of becoming a smart school can be carried out by a variety of individuals or groups in a school (or district). In this description of a smart-school process, that group will be referred to as the "improvement team."

School improvement involves learning new skills to solve complicated problems.

During stage 1, the improvement team prepares a school profile by gathering data on their current levels of effectiveness in meeting student performance goals. They identify discrepancies between performance levels and goal levels (or expected levels). Based on these discrepancies, the improvement team sets objectives for school improvement efforts. The following tasks define this stage:

Define important achievements for students in terms of measures. The first task is to operationalize important academic standards as actual measures. The improvement team might begin by raising these questions:

- What are the characteristics of the class, school, or set of schools targeted for an improvement plan?
- How well do stakeholders believe students are achieving the goals of the school?

- On what basis do stakeholders form their beliefs about student achievement?

One way to operationalize academic standards is through benchmarks. "Benchmarks" is a business term used to refer to describing the best results that currently are being obtained. Current performance at a school is judged in light of the best that could be expected. In a school setting, this means describing what children could be expected to do in various subject areas at various grades and reporting the findings. Reports would describe current student performance in relation to benchmark performance. One common way to operationalize expected performance is through the use of norms from published, standardized tests. However, these tests are subject to numerous shortcomings. To the extent possible, measures of academic performance should be chosen or developed that align with the benchmarks the improvement team identifies, so that the school's performance can be compared against the school's benchmarks.

Site councils must carefully focus on classroom targets such as teaching and staff development. It is <u>not</u> a site council's responsibility to micromanage a school.

Use the agreed-upon measures to construct a school profile. The measures should be administered to determine the current level of functioning of the students targeted for the improvement effort. The results will be the most important part of the school profile.

In addition to student performance, the following questions about approaches in current use can round out the profile:

- What is the match between the goals of the approaches used in the school and the school's curriculum goals?
- According to individual stakeholders, what are the strengths and weaknesses of the school's approaches?
- In what specific ways do the school's approaches accommodate the needs of diverse learners (special education, limited-English speaking, economically disadvantaged)?

Identify two or three improvement goals. After the improvement team identifies all the areas where school performance is below its desired level of performance, it recommends several improvement goals for each grade level or subject area. Grade-level or subject-area teams would then identify no more than two objectives for focusing school-reform efforts. At a later point in the process, the entire school staff and other interested stakeholders should participate in selecting the specific objectives to be the focus for school improvement efforts.

Stage 2: Defining the Scope of the Improvement Plan

Is the overriding goal to reform or to innovate? "Reform" refers to implementing procedures that have been demonstrated to work. "Innovate" refers to ideas that are untested. The scope of an improvement plan depends primarily on whether it is built around an innovation or a reform. Reforms merit widespread dissemination. Innovations require careful monitoring and evaluation, and a back-up plan to be put in place if students clearly are not benefiting from the innovation. The distinction between innovation and reform allows improvement proposals to be classified in one of two ways, according to their intent—to implement approaches that have been demonstrated to benefit students or to try out new ideas with the intent of benefiting students.

What is the extent of the improvement that is needed?

The greater the need, the more extensive the improvement will need to be. If actual student performance is low in all areas, then a more comprehensive change may be needed. If the actual student performance is low in only a few curricular areas, then change only in those more specific areas is warranted. If the discrepancy exists for only a certain population, then the school improvement objectives should focus on the needs of that population.

Not every innovation in education is an improvement. Only genuine improvements yield increases in student learning that are measurable.

Where will the approach come from?

An educational approach includes instructional tools and practices, professional development, and organizational support. There are three major sources for approaches; (a) create an approach using local expertise; (b) create an approach following a well-developed process, such as Sizer's Essential Schools, Comer's Social Development Model, or Levine's Accelerated Learning Model; or (c) adopt an already-designed, field-validated approach that focuses on teaching and learning, such as "Success for All" or "Direct Instruction."

An improvement team should begin the search process by identifying the people who will search for (a) the possible approaches for each targeted goal, and/or (b) identify a list of approaches that must be considered.

What scale of initiation and risk are acceptable—small, medium, or large (e.g., a classroom, school, several schools, or district wide)?

If a large-scale improvement is required, then the school must select approaches that (a) have proven successful in other wide-scale improvement efforts, or (b) have proven successful in a medium-scale implementation and include a plan for expansion that indicates there is a strong likelihood that the approach will be uniformly successful in a large-scale effort.

Site councils should have a plan for judging the potential value of an approach.

If a medium-sized adoption is required, then the approaches must have proven successful in a small-scale effort and the plan must make success a high probability.

If a small-scale adoption is required, the school may select approaches that appear to be promising and have informal assessments that can be administered at four- to six-week intervals to insure that students are learning.

Stage 3: Selecting Tools and Practices for the Improvement Plan

During stage 3, improvement teams search out and identify approaches that align with the improvement goals targeted by the school.

A major component of stage 3 is specifying the organizational support and professional development that must accompany the selected approach. There is no cookbook for successful educational reform; however, improvement teams should have a plan for judging the potential value of an approach. The following framework presents six important questions an improvement team can ask to evaluate both existing and proposed educational practices.

1. Are the Approach and Its Outcomes Clearly Defined?

If an approach and its measures are vaguely defined, it cannot be determined whether it has in fact been implemented, and consequently its effects are impossible to measure.

The first question your improvement team should ask when examining existing or proposed educational approaches is, "Are the approach and the intended student achievements clearly defined?" A clear description is easy to understand and comprehensive. "This innovative educational approach uses interactive holistic portfolios to enhance left-brain learning" sounds impressive but communicates little useful information.

A complete description of an educational approach provides a picture of what teachers and students will be doing, the instructional materials to be used, how teachers interact with students, and the amount of time devoted to instruction. Clearly defined approaches also convey what students will be able to do at the end of given periods of time and describe measures to evaluate whether these achievements take place. *These achievements should align with the site team's school improvement goals.* For example, a school's goal of all children reading simple stories by the end of first grade does not align with using a cooperative learning approach to get children to work together. If the school's goal had been increased cooperation, then that approach would align.

If an approach and its measures are vaguely defined, it cannot be determined whether it has in fact been implemented, and consequently its effects are impossible to measure. Measures of whether an approach has been implemented should focus on determining whether a match exists between what teachers do in the classroom and the description of the approach.

- Describe the approach in terms of philosophy and in terms of specific actions in the classroom.
- What are the intended results for teachers and for students?

2. What Evidence Exists the Approach is Effective?

People want to hear good news. "Adopt this program and your school will be transformed into a learning paradise." This claim may be true, or it may be wishful thinking. The point is that the effectiveness of an educational approach must be considered **before** your school decides to use it with students.

Be cautious about how the word "research" is used. Much "research" in education is not actually research but the opinion of people in the field.

You must be cautious about how the word "research" is used. Much "research" in education is not actually research but the opinion of people in the field. Research should be of high quality and ample to justify widespread adoption of an approach and the accompanying expenditure of funds and effort. For example, the success of an approach used in one classroom with a very talented teacher would not justify the use of this approach in a large number of classrooms with teachers of varying levels of performance. Instead, the approach would need to first be evaluated with more teachers.

If your improvement team proposes an approach, its effectiveness should be documented by answering questions such as these:

- If the approach has been implemented before:

Where were these implementations?

Did they involve teachers and students reasonably comparable to those found in the improvement site?

What were the results?

How do they compare with other approaches that have similar goals for similar populations?

In short, does the approach produce results as good or better than achieved by other established, successful approaches?

Do the results suggest that measures of student achievements will need to be adjusted? If so, in what ways?

- If the approach has not been implemented before because it is new:

Did the publisher follow a learner verification procedure in developing the approach? (Learner verification involves field-testing and revising the program based on student errors and on teacher difficulties in using the program.)

What were the results of the field-testing in terms of student achievement?

- If there are research studies on the approach:

Are explanations of the research clear enough for a layman to understand, yet detailed enough to be believable?

Are the results valid and substantial? If research shows only a slight gain in student learning in response to a large expenditure of money and effort, the value of the approach is questionable.

Accountability is not a process for identifying scapegoats.

In general, the education field suffers from a lack of research on significant issues. If research is not available to judge if an approach is effective, one option is benchmarking—visit schools using the proposed approach with representative students and teachers, observe their students' performance in the program, and review its evaluation results.

3. Is an Accountability Process Built into the Approach?

A lot of promising educational approaches fail because they can't help teachers determine whether they are successfully implementing the approach and if the students are learning at a desired rate and level of proficiency.

An accountability system should respond to these questions:

- Have student and teacher measures been identified or created that will contribute useful information about an approach's implementation?

In particular, are there benchmarks for student learning?

- Does the approach provide regularly scheduled observations for the purpose of supporting student and teacher performance in a timely and effective manner?

Is there necessary personnel for providing such support?

- Are assistance procedures specified for teachers and students who need additional support?

Are the procedures viable and effective?

4. Is the Approach Sustainable?

School improvement programs do not run themselves. They require substantial time and effort for a good payoff. It makes no sense for an improvement team to recommend an educational approach that cannot be sustained. Sustainability depends on how practical and reasonable it is to expect to implement and monitor an approach.

An approach that makes unrealistic demands on teachers' time is doomed to failure. For example, many reforms assume that teachers will be responsible for developing new curriculum materials. This can be an unrealistic assignment unless teachers are given adequate time for writing the lessons, trying them out with a range of students and teachers, accumulating data on student performance, noting problem areas, then rewriting the curriculum and trying it out again.

Sustainability depends on how practical and reasonable it is to expect to implement and monitor an approach.

An approach that makes unrealistic demands on teachers' time is doomed to failure.

Issues involving professional development are also critical. Helping teachers learn how to apply an educational approach should involve appropriate staff members for the same reason that baseball players aren't pulled off the bench and sent into hockey games: They may be skilled athletes, but their skills are the wrong kind. Advocates of innovative approaches often underestimate the amount of staff development needed to have all teachers reach a satisfactory level of implementation. Without adequate staff development and monitoring, the approach is likely to fail in many classrooms and ultimately be abandoned.

In judging, if an approach is sustainable, an improvement team should ask these questions:

- How much professional development is required for initial training? For follow-up training?
- Are the expectations for teachers reasonable in terms of the approach providing all or most of the required components?

In terms of expertise and compensation if teachers are required to create components?

- Are expectations reasonable for administrators?
- Are instruction, assessment, and professional development coordinated and sufficient?

How can local personnel develop or refine the expertise needed to maintain the implementation of the approach?

What kind and amount of outside support will be needed until local personnel take full responsibility for the implementation?

- Are the costs within the school's budget?
- If additional funds will be required, can they be specified and justified; also potential sources for those funds should be identified.

Approaches which are capable of bringing about significant improvement to the greatest number of children should be given priority.

Stage 4: Planning and Implementation

Once the school selects an approach, planning steps need to be taken for the implementation. Planning for the implementation steps usually requires six months or more.

Curriculum Materials. Doctors, cooks, mechanics, and carpenters all use tools to accomplish their work. So do educators. An often overlooked aspect of educational reform with important implications for learning are educational tools such as textbooks, computer software, and instructional videos. These tools are used in a range of settings and are often a focal point for classroom instruction. There is great variability among educational tools, which are usually marketed like other consumer products. The buyer must be cognizant that claims can sometimes be misleading or inaccurate.

Questions your improvement team can ask about the curriculum materials for a particular approach include:

- What curriculum material is needed for the teachers and students?
- What is the present cost of the material?
- What will be the cost in future years?
- When must materials be ordered so they are available for use when the implementation begins?

Monitoring Progress. The purpose of monitoring is to guide professional development, coordination, and reporting activities so that no students fall between the cracks. This can be accomplished by regularly monitoring the degree to which an educational approach is used well and the degree to which students are achieving at a satisfactory level.

There is great variability among educational tools, which are usually marketed like other consumer products. The buyer must be cognizant that claims can sometimes be misleading or inaccurate.

The purpose of monitoring is not to create anxiety. Monitoring should provide timely information for continuous improvement of the instructional program through professional development and coordination. A school must be able to determine if its agreed-upon reform efforts are being implemented appropriately. Schools and students benefit when a system exists to identify and assist individuals or groups not making adequate progress in the school.

Important questions an improvement team should ask about monitoring are:

- What type of periodic assessment will be used to monitor student performance?
- How much time is needed daily and throughout the year to monitor the implementation?
- Who will do the tasks needed to monitor the implementation?
- How much funding is needed to support a monitoring system?

Successful implementation of an approach is closely associated with staff development that involves everyone who will use the new tools and approaches.

Professional Development. School improvement often requires people to change the established order and comfortable way of doing things. Thus, the more successful teachers initially are with a new approach, the more willing they are to become active supporters of it. Successful implementation of an approach is closely associated with professional development that involves everyone who will use the new approach.

An important part of professional development is learning about teaching tools and practices used in an approach. Perhaps as important, the rationale and expected student outcomes for using an approach is a significant factor in the motivation of staff members. Staff development should also include practice teaching lessons that will come up in the near future. This will increase the probability that teachers and students will be successful.
Coaching, on the other hand, may involve a staff developer or "lead" teacher working in a classroom, modeling how to work with students, observing and giving the teacher supportive information to improve his or her use of an educational approach.

In talking about how successful businesses support their employees, Albert Shanker, president of the American Federation of Teachers, said this about GM's Saturn plant:

> How did Saturn find these smart, flexible, and disciplined workers? It didn't find them; it used an impressive training program to give workers from 136 other General Motors plants the information and skills and ongoing help they needed to participate in this new way of running an automobile plant.
>
> The original team members received more than 400 hours of training within their first few months at Saturn, and even now, new employees take part in a kind of internship. During the first two or three months, they split their time between classroom and on-the-job training.

In a related vein, several of Demming's 14 points of total quality management imply intensive support for teachers for learning new, effective practices: (3) build in quality; (5) improve constantly and forever; (6) include modern methods of training on the job; (10) eliminate slogans; (13) target education and self improvement.

A proposal for a professional development plan should answer the following questions:

- Who will conduct the professional development? If outside professional developers are needed, consider all costs for hiring them. Also, consider how long professional development will be necessary. Will on-site people be able to provide professional development in the short and/or long run?
- How much professional development will be necessary? When will inservices take place? What extra costs are involved to release staff?
- How much in-classroom professional development is necessary? Who will do it? What are the costs?

Coordination of Efforts. Educational reform does not occur in a vacuum. This means your school's existing programs and schedules will need to be coordinated to minimize interference with the implementation of the approach.

Programs and schedules will need to be coordinated to minimize interference with the implementation of the approach.

Instructional approaches and priorities for programs such as TAG, Chapter 1, and special education should be coordinated to avoid conflicting approaches which may confuse students and communicate conflicting messages to teachers.

Adequate teaching time should be provided for implementing the approach. The amount of time allocated should take into consideration that students with greater deficits in knowledge may require more time to reach desired levels of performance.

When planning how to coordinate a school's program, your improvement team's plan should answer the following questions:

- How much time is needed to implement the program? How will that time be found in the schedule?
- Are teachers receiving any conflicting messages and demands?

Accountability Reports. Once an implementation is underway, the school's administration should provide its stakeholders—students, parents, staff, and community members—with regular status reports. The reports should focus on children's performance and on what is being done to help the children who are not at the desired level. Agreement needs to be reached on what the reports will contain, when they will be produced, to whom they will be given, and what actions will need to be made.

Program Evaluation. A school will want to celebrate its accomplishments while always working to refine its school improvement efforts. Neither activity can occur without an evaluation plan to determine the overall effectiveness of the project in improving student learning. An evaluation plan must be developed *before* the beginning of the implementation, particularly to make sure that student performance data are gathered before the implementation begins. This means that information needs to be obtained to determine the children's level at the beginning of the implementation so it can be compared with the children's performance at the end of each year.

The site council must develop an evaluation plan before the beginning of an implementation.

Evaluation is complex. For example, if the evaluation is going to compare the performance of children in the new project with children in other approaches, care should be taken to be sure the two situations are comparable. Some schools have high student turnover rates, and it's not uncommon for students who begin first grade in one school to attend sixth grade at a different school. Comparing the sixth-grade performance of a school with high turnover to the sixth-grade performance of a school with low student turnover would not provide an accurate comparison of the programs.

Some important questions about evaluation your improvement team should ask are:

- What assessment measures will be used?
- Do the measures reflect what has been taught?
- Will a comparison be made with other approaches?
- Who will the comparison group be?
- What funds are necessary for the evaluation?

For major reform efforts, it's advisable to seek assistance from an evaluation specialist.

An Abbreviated Decision-making Process for Creating Smart Schools

Figure 1 illustrates a decision-making flowchart that can guide schools in improving teaching and learning. Based on the school profile, the improvement team and eventually the entire school staff would be involved during stage 1 in determining areas where they were not satisfied with student performance (the first decision diamond).

The first step in defining the scope of the improvement plan (stage 2) is to determine if benchmark (or better) performance has been achieved elsewhere (the second decision diamond). Innovative approaches (approaches that are completely new and have not yet been tested or validated) involve a much greater risk than using validated approaches. The risk in using an innovative approach is only justified if a school is (a) already effectively using the state-of-the-art in validated approaches or is achieving the benchmark levels

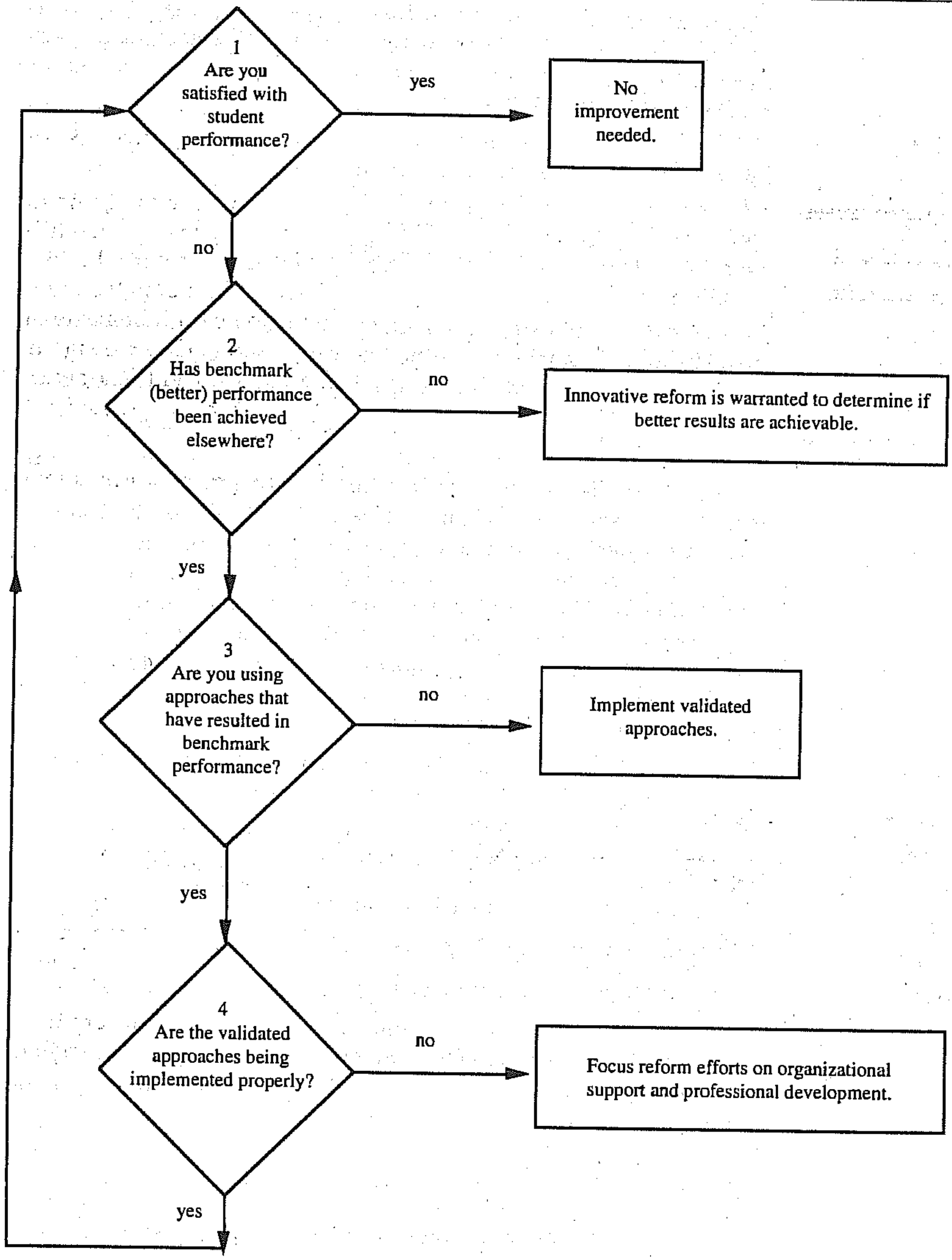

Figure 1. The decision-making process that can guide schools in determining some important features of the tools they should be looking for.

of performance, and (b) the school is still not satisfied with the results. Only if the answer to this question (the second decision diamond) is "no" is a school justified in selecting an innovative approach for implementation.

Think of this handbook as a road map for your site council.

As noted in the discussion of stage 2, innovative programs can range in extensiveness (one curricular area for one group of students to a comprehensive new approach for all students) and can be created locally or adopted from elsewhere. However, innovative programs entail less risk if they are first piloted with a relatively small number of students.

If a school is not using validated approaches, then the selection of new, validated approaches is the first order of business. This decision point is illustrated in the third decision diamond in Figure 1.

If validated approaches are in use, but benchmark performance is not being accomplished, the factors of organizational support and professional development that influence the effective use of the approaches will need to change. This decision point is illustrated in the fourth decision diamond in Figure 1. The problem may lie exclusively in issues of organizational support and professional development. The school may not be implementing the approaches appropriately. In this case, the school reform plan can center on changing aspects of the organizational support and professional development so that consequent benchmark results are achieved, without selecting a new approach.

Many aspects of professional development and organizational support have been identified in the body of school-effectiveness research. Some of the factors are:

- The school is characterized by a business-like atmosphere.
- Student achievement is celebrated.
- Students are properly placed.
- Academic learning time is high.

Approaches which lower expectations for students, or delay academic work in favor of non-academic activities are not school improvement.

In summary, effective school improvement may warrant a change in approach; in other cases, it may warrant a change exclusively in aspects of organizational support and professional development, without selecting new approaches.

Conclusion

The changing demographics of our population and the increasingly competitive international economy are putting increased pressure on our education system to become more effective. The structures that were functional during the 1950s are clearly no longer functional in the 1990s.

This urgency might explain why educational reformers are obsessed with having answers—restructured schools, national standards, choice, etc. Having answers is not actually a problem. In fact, we have too many answers. There have been over 350 national panels, commissions, and committees in just the past ten years. The problem is recognizing and acting on a good answer.

People inside and outside of education often accept an innovation's attempt to promote change as a substitute for change itself, or they assume that putting in an innovation is synonymous with improvements in student learning. They declare projects a success because teachers are involved in decision-making, use hands-on teaching and/or cooperative learning strategies, and so forth. But they often don't stay with the project long enough to determine whether learning improved in a substantial fashion.

A central mission of the National Center to Improve the Tools of Educators (NCITE) is to contribute to the development of smart schools. A smart school uses information about student learning to set goals and seek out, select, and implement productive answers that align with their goals. The smart school continues to monitor learning to keep the improvement cycle going. Smart schools look for productive answers that can result in substantial achievements for students. These answers involve measures of student learning, instructional tools and practices, professional development, organizational support and ties to the community. NCITE's purpose is not to endorse a particular answer but to mobilize and inform groups that can lend expertise and motivation for schools to become smart. These groups include teachers' unions, businesses, school boards, community groups, policymakers, professional organizations, and publishers. The contribution each group can make is described in the following publications:

Creating Smart Schools: The Contribution of Schools Boards

Creating Smart Schools: The Contribution of Teachers Unions

Creating Smart Schools: The Contribution of Community Organizing Groups

Creating Smart Schools: The Contribution of Businesses and Foundations

Creating Smart Schools: The Contribution of the Profession—Lessons from Health and Engineering

As a road map, this handbook does not dictate either destinations or routes, but it does help the travelers organize the trip and understand when important decisions are to be made and what the implications of those decisions might be.

Bon Voyage

Summary of A Handbook for Creating Smart Schools

Stage 1. Setting Improvement Goals

1. Define important achievements for students in terms of measures.
2. Use the agreed-upon measures to construct a school profile.
3. Identify two or three improvement goals.

Stage 2: Defining the Scope of the Improvement Plan

1. What is the extent of the improvement that is needed?
2. Where will the approach come from?
3. What scale of initiation and risk are acceptable—small, medium, or large (e.g., a classroom, school, several schools, or district wide)?

Stage 3: Selecting Tools and Practices for the Improvement Plan

1. Are the Approach and Its Outcomes Clearly Defined?
2. What Evidence Exists the Approach is Effective?

- If the approach has been implemented before:

Where were these implementations?

Did they involve teachers and students reasonably comparable to those found in the improvement site?

What were the results?

- If the approach has not been implemented before because it is new:

Did the publisher follow a learner verification procedure in developing the approach? (Learner verification involves field-testing and revising the program based on student errors and on teacher difficulties in using the program.)

What were the results of the field-testing in terms of student achievement?

- If there are research studies on the approach:

Are explanations of the research clear enough for a layman to understand, yet detailed enough to be believable?

Are the results valid and substantial? If research shows only a slight gain in student learning in response to a large expenditure of money and effort, the value of the approach is questionable.

3. Is an Accountability Process Built into the Approach?

- Have student and teacher measures been identified or created that will contribute useful information about an approach's implementation?
- Does the approach provide regularly scheduled observations for the purpose of supporting student and teacher performance in a timely and effective manner?

- Are assistance procedures specified for teachers and students who need additional support?

4. Is the Approach Sustainable?

- How much professional development is required for initial training? For follow-up training?

In terms of expertise and compensation if teachers are required to create components?

- Are expectations reasonable for administrators?
- Are instruction, assessment, and professional development coordinated and sufficient?
- Has the success of the professional development been documented in terms of the ability of staff members to perform at uniformly high levels following the proposed staff development?

5. Is the Approach Equitable?
6. Are the Costs of the Approach and its Implementation Reasonable?

Stage 4: Planning and Implementation.

1. Curriculum Materials.
2. Monitoring Progress.
3. Professional Development.
4. Coordination of Efforts.
5. Accountability Reports.
6. Program Evaluation.

References

Alessi, G. (1988). Diagnosis diagnosed: A systemic reaction. *Professional School Psychology, 3*(2), 145-151.

Business Roundtable (1988). *Essential components of a successful education system: The Business Roundtable education public policy agenda.* Washington, DC: The Business Roundtable.

Fuhrman, S., & Elmore, R. (1992). *Takeover and deregulation: Working models of new state and local regulatory relationships.* New Brunswick: The State University of NJ.

Glickman, C. (Sept., 1992). The essence of school renewal: The prose has begun. *Educational Leadership, 50*(1), 24-27.

Kearns, D., & Doyle, D. (1991). *Winning the brain race—A bold plan to make our schools competitive.* San Francisco: ICS Press.

National Alliance of Business (1993). *Cutting Edge of Common Sense in Brief.* Washington, DC: National Alliance of Business, Information Services.

Rigden, D. W. (1991). *Business/School Partnerships.* New York: Council for Aid to Education.

Rigden, D. W. (1990). *School restructuring: Designing schools for better student learning.* New York: Council for Aid to Education.

Seymour, Daniel T. (November, 1991). TQM on Campus: What the pioneers are finding. *AAHE Bulletin*, 10-13, 18.

Stone, N. (1991, March-April). Does business have any business in education? *Harvard Business Review*, 46-61.

The Mattawan, Michigan Model

Key Features of a Working School Improvement Model for Schools

For more information see "Learning is Our Business" by Howard Farris (Mattawan Bd President), Doug Carnine, and Jerry Silbert in The American School Board Journal, *December, 1993, pages 31-33.*

Key personnel and duties:

Instructional facilitator

- Helps teachers create and administer tests which students take every quarter.
- Helps analyze the results of the tests and implement the solutions to meet the needs of every student, not just the high and low performers.
- Develops the skills of teachers, especially new teachers, in a non-threatening atmosphere of growth. Does not formally evaluate teachers for contract renewal decisions.
- Ensures that each *student* and each instructional *group* is progressing at an optimal rate.

School Board Student Learning Subcommittee

- Composed of 3 school board members or appointed parent representatives.
- Reviews reports on student learning.

Key reports to the School Board Subcommittee:

Quarterly student performance reports

- Each group's rate of progress is determined by comparing a projected goal with the group's actual progress: If a group's performance falls short of the goal, the instructional facilitator and teacher develop a corrective course called an action plan.
- If an individual student's performance level indicates that he or she is having difficulty in a subject, the action plan might also call for the instructional facilitator to initiate diagnostic testing, including classroom observation, and work with the teacher to give the child special help.

Action plans

- Action plans propose solutions, either using the available resources or using new resources (e.g., an additional teacher to reduce class size). When significant new resources are required, the action plan is submitted to the full board for approval. The plan states the problem in terms of student performance and provides a rationale for the proposed solution and related expenditure.
- For groups of students, solutions might involve changing the teaching practices or the curriculum.
- For individual students, solutions might include extra tutoring or assistance from a specialized teacher.

Procedures for new adoptions as solutions

- District officials search continuously for successful programs that lead to high levels of student performance on academic outcomes.
- Step one: Obtain data from other elementary, middle, or high schools.
- Step two: Analyze the data to identify high performing groups of students from backgrounds similar to the school population.
- Step three: Consider adopting the practices and programs being used in the high-performing schools.

Working Model: Mattawan, Michigan

- Ranks among the lowest (517th out of 524) in per-pupil expenditure in Michigan.
- Consistently ranks among the state's best on statewide assessments.

A Teacher's Guide for Reading Research

Bonnie Grossen
University of Oregon
Barbara Ruggles and Sandy Bailie
American Federation of Teachers

GUIDES

The primary question facing teachers, and the ultimate question facing educational researchers, is "What can we (I) do to improve student learning?" The answers to this question comprise the knowledge base of teaching procedures and tools that empower teachers to maximize student learning. The purpose of this guide is to help teachers become good consumers of research and gain access to this knowledge base.

Research, Craft Knowledge, and Dogma

Research and craft knowledge involve scientific thinking. Dogma does not. Scientific thinking uses logical reasoning to derive knowledge from evidence. Scientific thinking can be both formal and informal. "Research" describes formal scientific thinking and "craft knowledge" describes informal scientific thinking. "Dogma" describes beliefs that derive neither from formal (research) nor informal scientific thinking (craft knowledge), but seem to come from some authority. The following section explains the significance and meaning of these three concepts in greater detail.

Research that answers the question, "What can I do to improve learning," involves formally controlling experimental situations to test hypotheses against empirical evidence (objective observations of reality). The experimenter manipulates the teaching practices (i.e., the "independent" variable), controlling all other variables so they do not affect the results, and observing the results in terms of student learning (i.e., the "dependent" variable). Experimental research that evaluates how student learning "depends" on something the teacher can "independently" control is the most reliable way to build a knowledge base about what teachers could do to improve learning. Table 1 provides a simple description of such a scientific experiment.

A formal science derived through rigorous experimental studies, such as described in Table 1, is not born overnight, but continues to develop over time. Therefore, research cannot describe everything a teacher does. A teacher can and should integrate his or her craft knowledge (a science of teaching developed more informally from classroom experience) with the knowledge derived from research (a science of teaching developed from carefully controlled experimental studies) in a practical scientific model of teaching.

Craft knowledge describes the knowledge teachers gain by thinking scientifically about their own experience. Similar to research, craft knowledge derives from closely observing student behavior to evaluate teaching practice. For example, "don't smile for the first six months" is a piece of craft knowledge that experienced teachers often pass on to new teachers. As far as we know, there is no research testing the truth of this statement. It is probably the result of many teachers' reflecting on their first year of experience and noting that they were too familiar with their students. The ability to express respect and kindness to students and yet remain the person in charge of a large group of energetic youth does not come easy to many new

Table 1. A simple description of an ideal scientific investigation.

Hypothesis: Problem-solving toolbox A will help students solve math problems.

Independent variable: Problem-solving toolbox A.

Dependent variable: A set of math problems for students to solve.

Design: A sample of students are assigned to two instructional groups. The two groups of students are equivalent in ability and background knowledge and in everything relevant to mathematical problem solving before the experiment begins. During the experiment, the instruction to both groups is equivalent in every regard, except for the fact that one group uses the problem-solving toolbox, the other group does not. When the instruction is completed, both groups take the posttest. The performance of the two groups is compared to see if the group using the toolbox achieves higher scores.

teachers. This piece of craft knowledge serves to caution new teachers about the dangers of becoming too friendly with students.

Teachers' craft knowledge should not contradict scientific research, but rather fill in the gaps of scientific research. Any contradiction that occurs must be theoretically resolved so that the teacher's practices remain consistent with the scientific data. Either there were problems with the design or interpretation of the data in the study, or the teacher's practices may need to change. Findings from carefully controlled studies such as that in Table 1 provide more reliable knowledge than teacher's craft knowledge, derived through informal study and observation. Do now Activity 1 below to evaluate the significance of the concepts research, craft knowledge, and dogma.

The answers to part I represent craft knowledge. If the answers to part II were *a*, they represent dogma. Research indicates that both answers are *b* (see Appendix A). In this case what people are most likely to hear is true about brainstorming contradicts actual research. Robert Weisberg (1986) summarizes the research on brainstorming and problems solving as follows:

> "If one wishes to solve a problem effectively, one should try to determine as precisely as possible what criteria the solution must meet before starting work on the problem, try to keep these criteria in mind as one works, and work alone" (p. 66).

Weisberg also explains how a very effective marketing program smothered these research findings and persuaded American education and education abroad that brainstorming in groups without criteria was more effective. Beliefs that are promoted by some authority, such as these popular beliefs about brainstorming were, yet contradict scientific evidence, are dogma.

Which type of knowledge has governed your teaching behavior? Especially teachers who found in Activity 1 that their craft knowledge did not agree

Activity 1: Comparing research, craft knowledge, and dogma

Imagine these kinds of problems for your students to solve:

(a) "How can Americans have more influence in foreign policy?"

(b) "Anticipate consequences that might arise if people were born with an extra thumb on each hand."

(c) "Increase the number of European tourists visiting the U.S."

(d) "Invent brand names for a cigarette, a deodorant, and automobile."

I. Answer these questions based on your own personal craft knowledge (your experience):

—— 1. Is it more efficient to (a) have students work in groups to brainstorm?
(b) have students work individually to brainstorm?

—— 2. Is it more efficient to (a) have students brainstorm first, then develop criteria for evaluating a solution?
(b) have students develop the criteria first, then brainstorm?

II. Now answer these questions based on what you have been told is research or best practice.

—— 1. Is it more efficient to (a) have students work in groups to brainstorm?
(b) have students work individually to brainstorm?

—— 2. Is it more efficient to (a) have students brainstorm first, then develop criteria for evaluating a solution?
(b) have students develop the criteria first, then brainstorm?

Discussion:

Do you find there is a contradiction in your answers to part I and II?
If so, how would you resolve the contradiction?
See Appendix A to find what research says is best practice.

with the dogma, but did agree with research that they were unaware of, may also acknowledge that their teaching behavior has been governed by dogma, rather than their craft knowledge. For these teachers their craft knowledge, derived from their own personal experience, would have been more reliable than listening to some authority promoting dogma. Teachers do not intentionally let dogma govern their behavior, but unfortunately, dogma is often touted as research. Teachers innocently follow dogma and contradict research everyday in the classroom. Following dogma likely will cause learning to deteriorate.

Teachers' craft knowledge should not contradict scientific research, but rather fill in the gaps of scientific research.

Scientific thinking (the use of research and craft knowledge) has increased the productivity of agriculture, medicine, and other professions. Scientific thinking should also increase the ability of teachers to improve learning. A good parallel model for scientific thinking is agriculture. (See Worrall & Carnine in this issue for a comparison of education with medicine and engineering.) Farmers evaluate the effectiveness of their farming practices by looking at the size of their harvest. Their goal is to get higher yields by adjusting and changing their farming practices. Similarly, scientific teachers evaluate the effectiveness of their teaching practices by looking at the levels of student performance. Their goal is to get higher levels of student performance (dependent variables) by adjusting and changing teaching practices (independent variables).

Individual farmers could conduct their own scientific research by systematically varying farming practices to see how these changes affect the harvest. However, it is preferable to have agricultural researchers take the risks and costs involved in carefully controlling variables and studying their effects. This research information helps to build a shared knowledge base about farming. Similarly, it is more efficient for teachers to use educational research than to conduct their own individual scientific research by varying teaching practices to see how these variations affect student learning.

When farmers apply research, they do not ignore the size of their harvest and assume that applying research will maximize their harvest yield. For example, learning that irrigation improves growth of potatoes in Idaho, Idaho farmers would not simply irrigate and ignore the fact that their potatoes rotted from overwatering. Farmers would apply the research and thoughtfully evaluate it, adding their own craft knowledge regarding the amount of water that seems optimum for their particular fields, for example. Farmers never stop looking at their own yield as they apply research. In other words, farmers apply agricultural research in the context of developing their craft knowledge. Craft knowledge is not personal beliefs though, without regard for the evidence available in one's personal experience. Craft knowledge derives from less formal "research," which involves comparing this year's yield with last year's yield, comparing the current yield with the average yield or comparing it with the best yield. Craft knowledge is based on evidence, unlike dogma which is based on beliefs without regard for evidence.

Similar to farmers, teachers also should not simply apply educational research and trust that student performance will be maximized. Teachers should continue to develop their craft knowledge using less formal research and scientific thinking, which involves comparing the performance of this year's group of students with that of last year's, comparing their current levels of performance with the average performance (norm-referenced tests) or comparing them with the best performance achievable (benchmark criteria on criterion-referenced tests). The great danger in applying what one is told is research and ignoring the actual results, is that dogma is always dressed as research. In today's climate of sophisticated marketing tactics, it is always better to be skeptical than trusting.

Scientific thinking (the use of research and craft knowledge) has increased the productivity of agriculture, medicine, and other professions. Scientific thinking should also increase the ability of teachers to improve learning.

The Anti-Science Views

An anti-science attitude encourages education to follow dogma. A common criticism of science is that it is not perfect, observations are never objective, and so on; therefore, the scientific method is not a reliable way to build knowledge. Of course, science is not perfect. The experiment described in Table 1 sounds ideal. In reality the experiment could never be carried out as perfectly as it is described in Table 1. Reality is not as perfect as theory. However, knowledge gained using studies that are closer to the ideal (i.e., more rigorous and formal), is more reliable than knowledge gained using less formal quicker access to a larger body of research than original studies do. The pitfalls to avoid in reading both types of research are described later.

Reviewing Summaries of Research

Research syntheses are very valuable because they allow the reader to efficiently survey a lot of research. It is very important though to look for a "scientific style" in the writing. John T. Bruer (1993) in *Schools for Thought* exemplifies a scientific style. Bruer helpfully summarizes relevant research to support his points and conclusions. The data and logic behind Bruer's thinking are clearly described so that the reader may review all the evidence before accepting the conclusions Bruer makes. The references are clearly cited so that the reader may go to those original sources and review the research data in detail, to see if Bruer omitted something significant or slanted the results in some way.

Sometimes, in general literature about teaching, conclusive statements are made without a summary of the research. This is also acceptable *as long as research citations of experimental studies* are provided for the critical reader. The reader should beware though. Often citations are provided that do not include any data but rather only cite other opinions. The following parody modified slightly from Sprick (1992) illustrates this problem:

> Many people assume that if they see printed material in a book or journal that looks like this, they are examining research (Grossen, 1992). What they do not realize is that professional opinion can be referenced in the same manner (Ruggles, 1993).
>
> A small group of prolific professionals with strong beliefs can write a great deal and quote each other's ideas (Bailie, 1994; Grossen, 1993). They can create a circular research base that may appear to be research (Bailie, 1995), but may, in fact, just be bullshit (Engelmann, 1965).

A true skeptic will always go to the original research to verify the assertions made in summary sources, especially if no description of the research study is provided.

Reviewing the Actual Research Studies

Reading studies requires checking both the data analysis and the soundness of the reasoning from those data. Checking the data analysis may sometimes require a lot of technical knowledge about statistics and so on. Checking the soundness of the reasoning requires sound logical thinking. The greatest misuse of research seems to come from the latter, unsound reasoning from data, rather than the former. So in this guide we will emphasize evaluating the soundness of the reasoning and assume the data analysis is proper and accurate. Taking this approach, we will gain a lot of ground in becoming good consumers of research without having to learn a lot about statistics.

In answering the question "What can we do to improve student learning?" your goal is to find procedures / tools that you can use in your classroom. Not only must these procedures / tools be sufficiently described that you understand exactly what to do to implement them, or know where to get the specific knowledge to implement them, but the results must be worthwhile. The students in the research study must be performing at a level that is higher than what your students currently achieve.

Below are the steps in reviewing a research study. Each step will be explained by illustrating how it is applied to the theoretical study in Table 1. A summary of these steps can be found in Appendix B.

Step 1. What is the research question? The first thing to do in reviewing scientific research is to look for the research question. The research question in the problem-solving study was, "Does problem-solving toolbox A improve students' mathematical problem solving?" The question should state both the dependent variable (students' mathematical problem solving) and the independent variable (problem-solving toolbox A).

Step 2. Look for the task or test that students were asked to do. Does the test, or does the sample of items from the test, represent what you understand of the words in the research question that describe the dependent variable, to mean?

Assume the test in our theoretical study contained items like these: 2 + 4 = __ ; 79 X 231 = ___. Do these items fit your definition of "mathematical problem solving"? If not, revise the research question in your mind. For example, a better way to describe the dependent variable might be "mathematical computation," rather than problem solving. In research, we call this match between the actual test or task and what the researcher claims to be evaluating "validity." Validity is determined by a consensus of opinion. If most people would agree that the above two items represent problem solving, then the test would be a valid measure of problem solving. But if most people think of something else as problem solving, then the validity of the test is in question.

Just because the actual test or task does not match your understanding of "problem solving" does not mean that the study has no merit though. It just means that the results may not be interpreted to mean what you might have thought they would mean, if you had not looked at the actual test or task that was used to evaluate student performance. You may interpret the data your way, as long as you are logical and your interpretation is consistent with other research.

When you read research, be sure to look at the data as if you did not know what the researchers' conclusions were.

When you report to students and parents regarding the quality of a students' work, a main concern of all parties is that the reports are fair. In research terms, fairness (consistency) is called "reliability" and it can be quantified. An acceptable level of reliability is a "reliability coefficient" of at least .70. (The possible range is 0 to 1.0.)

Step 3. Look for the data. If you were a teacher conducting the problem-solving toolbox study, you would not start making any conclusions without looking at students' problem solving scores. When you read research, be sure to look at the data as if you did not know what the researchers' conclusions were. Interpret the actual data on student performance for yourself. Compare your independent interpretation with the researchers'.

In interpreting data, you look for differences in mean scores. If one instructional treatment is better than the other, we would expect the mean score for the group who received that treatment to be better. However, chance is always a possible explanation for a difference in mean scores. In fact, the likelihood that two groups of children will score exactly the same is very small so the likelihood that a small difference in means occurred by chance is very great. Research studies use statistics to rule out chance. Without ruling out chance, you cannot really conclude that just because there is a difference in the means, one group was better than the other. When the difference in the means is great enough to rule out chance as an explanation for the difference, research studies usually report that the differences were "significant." To describe the differences as "significant," the researcher must apply an acceptable statistical test of significance and find that the probability that the results could have occurred by chance was less than 5 in 100 (i.e., $p < .05$).

If there are no data cited or reported, the study is not research and is probably no better than your own personal opinion.

After you have looked at the data and come to your own conclusions, check if the researcher's conclusions agree with yours. If there is a discrepancy between the researchers' conclusions and your own, review the data again to see if you missed something.

Step 4. Review the procedures. The procedures section of a research study should describe the independent variables in detail. Read the details of the procedures to determine what the variables were. Identify the critical differences in the procedures that might explain the results. What alternative procedures or tools were not compared which might be just as effective, or possibly more effective? Read with enough care that you actually learn what you could do in your classroom to improve learning.

If there is a discrepancy between the data and your personal craft knowledge, review the design and experimental procedures to see if the experiment was fair and if the procedures actually align with what the experimenter said was done. If, for example, the problem-solving toolbox in the study described in Table 1 involved nothing more than a number of drill sheets for number facts, you might note that "problem-solving toolbox" does not accurately describe the procedures.

Step 5. Compare with your craft knowledge and your teaching situation. If you are unable to resolve the discrepancy between the data and your craft knowledge through more careful reading, then try the procedure that seems to contradict your beliefs in your classroom and take your own careful data on the results. You may need to revise your craft knowledge to accommodate new scientific knowledge.

Practice Activities: Identifying Pitfalls

The No-data Problem

Reference material:

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Washington, DC: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics.* Washington DC: Author.

The lack of any data supporting a set of teaching recommendations is currently an all-too-common problem. The highly praised teaching procedures recommended by the National Council of Teachers of Mathematics (1989, 1991), for example, are not supported by data. (Note: The recommended mathematics goals are distinct from the recommended teaching methods.) The documents claim only to be a research agenda, not a research-based document. One document states that "one reviewer of the Working Draft of the Standards suggested the establishment of some pilot school mathematics programs based on these *Standards* ..." (p. 253, 1989).

Structural analysis of responses to "Saturday Shopping"

Problem Structure: ? − [b − d] = 4.43

Solution Structure[1]	Dyad: 4C[2]	4M	5Mi	5C	5M	6Mi	6C	6M	7Mi	7C	7M
[b + d] + 4.43								X			X
[$b + d_1$] + 4.43	X		X								
[$b + d_2$] + 4.43					X						
[b] + 4.43							X			X	
→ [b − d] + 4.43											
[b + d] 4.43									X		
[b + d] (−4.43)				X[3]							
[b + d] − 4.43		X									
[b] 4.43					X						

[1] d_1 and d_2 refer to 1.50 and .25 respectively
[2] C—control, M—maximum, Mi—minimum
[3] Stated in protocol that answer is [b + d] but subtracted 4.43 from [b + d] on planning board.

Not even this pilot evaluation occurred before the NCTM teaching procedures began to be adopted nationwide. Though no research is cited in the NCTM documents to support the efficacy of the NCTM's recommended practices, educators across the nation seem to accept them as if they were a proven reform. Hopefully, luck is with the nation and the opinions of the members of the NCTM committee represent an improvement in teaching practice. It is, however, a very big risk, one that farmers would not take in comparable proportions with their crops.

Any document, article, or position paper recommending any teaching practices should provide supporting research. Teachers may choose to use the recommended practices, but they should always bear in mind the hypothetical nature of the recommendations, and carefully evaluate the performance of their students in the classroom. Any document used to force teachers to adopt new practices should be supported by scientific evidence.

The Irrelevant-Data Problem

Reference material:

Schoenfeld, A. H. (1989). Teaching mathematical thinking and problem solving. In L. B. Resnick & L. E. Klopfer (Eds.), Toward the thinking curriculum: Current cognitive research (pp. 83-103). Alexandria, VA: Association of Supervision and Curriculum Development.

Schoenfeld (1989) reports a study where he taught mathematical problem solving by circulating around the classroom asking three questions: (a) "What (exactly) are you doing? (Can you describe it precisely?)," (b) "Why are you doing it? (How does it fit into the solution?)," and (c) "How does it help you? (What will you do with the outcome when you obtain it?)." Schoenfeld reports the effects of this intervention as follows:

> Students are asked these questions early in the term. They are generally at a loss to answer them (and I encounter a significant amount of hostility and resistance). When the students realize that the questions will continue, they begin to defend themselves by discussing the answers in advance. By the end of the term, discussing the questions has become habitual for them. (p. 98)

The only data Schoenfeld reports do not show an increase in success in problem solving. Rather these data describe how one pair of instructed students in Schoenfeld's University class spent their time in solving a problem. Schoenfeld does not describe how he selected this pair of students, whether they were randomly selected, or were perhaps the best example of what he hoped to achieve. Nor does he describe how the other groups' problem solving behavior compared with this pair's.

Any document, article, or position paper recommending any teaching practices should provide supporting research.

Schoenfeld presents no data on the effect of this intervention on problem solving success. Successful problem solving does not even seem to be Schoenfeld's goal any more. He states that "the point here is not that the students managed to solve the problem, for to a significant degree solving non-standard problems is a matter of luck and prior

The Obvious-Experimenter-Bias Problem

Reference material:
Vance, J. H., & Kieren, T. E. (1972). Mathematics laboratories—more than fun? School Science and Mathematics, 617-32.

Vance and Kieren compared three instructional treatments, a mathematics laboratory group, where students used manipulatives to discover concepts with partners, a class discovery group, where the teacher used the manipulatives to present the discovery lesson to the whole class, and a control group, where the usual mathematics instruction was used. Means and standard deviations are not reported, as is conventionally expected in a scientific study. Nevertheless, the experimenters summarize the data to some extent for the reader.

One of the measures was a daily test over the content of the lesson. "Analysis of the total scores obtained by students on the ten review sheets indicated no significant difference between learning in the laboratory and class discovery settings except for average and low ability seventh graders. For these samples the class discovery group scored significantly higher than the lab group. But the Grade 8 students and high ability Grade 7 students who worked in pairs from written instructions appeared to have learned as much as those in the teacher-directed class setting" (p. 620).

Another measure was a cumulative test of problem solving. "Students from the experimental groups performed significantly better than control students. Group mean scores of the class discovery students were generally slightly higher than those of the lab students, although differences were generally not significant" (p. 621).

A measure of divergent thinking was also administered. "The mean number of acceptable responses was highest for the class discovery group, followed by the lab group, then the control group." The level of significance is not reported, though convention would dictate that these statements would not be made unless the differences were significant.

On all the above measures, the teacher-directed class discovery group scored higher, either significantly or nonsignificantly. Only on attitude measures did the trend shift to favor the lab treatment. "On the three subscales of the instrument the lab group rated highest, followed by the class discovery group and the control group. In particular, the lab students rated highest in feeling that learning mathematics is fun or enjoyable and in the view that mathematics is a subject which can be explored experimentally." And on another reaction measure: "Although students from both experimental groups reacted positively to their respective program, the reaction of the lab students was more highly favorable than that of the class discovery students. While both groups registered uncertainty about what content they had learned from new experiences, a significantly larger number of students in the laboratory felt they were gaining something from the lab than did students in the Class Discovery setting" (p. 623).

Besides lacking rigor in the reporting of data so that others may check and review the results, the Vance and Kieren study presents conclusions that misrepresent the results of the study.

Yet in the conclusions, Vance and Kieren seem to imply that the laboratory group performed better than the class discovery group: "Although, in terms of new learning, a teacher demonstration-and-discussion approach [class discovery] appears to be *at least as effective* [emphasis added] as a manipulative laboratory with the concrete materials with students at these grade levels, students strongly prefer the laboratory situation because of the freedom it provides for experimentation and for working independently of a teacher" (p. 623).

Besides lacking rigor in the reporting of data so that others may check and review the results, the Vance and Kieren study presents conclusions that misrepresent the results of the study. The two experimental treatments seemed like a fair comparison, but the less effective treatment was represented as the better treatment.[1]

The Wrong-Descriptor Problem

Reference material:
Bay, M., Staver, J., Bryan, T., & Hale, J. (1992). Science instruction for mildly handicapped: Direct instruction versus discovery teaching. Journal of Research in Science Teaching, 29(6), 555-570.

Bay, Staver, Bryan, & Hale compare activity-based learning with teacher-directed instruction using students with disabilities. They conclude that the results of their study contradict previous research that has found teacher-directed instruction more effective with special education subjects.

However, the instruction that Bay et al. call teacher-directed is not the same direct instruction supported by previous research. Previous research supports instruction that is highly interactive with the teacher. The intervention that Bay et al. call "activity-based" seems to align more closely with research-based direct instruction. Bay et al. report that the number of teacher-initiated teacher-student interactions was higher for the activity-based instruction (412 teacher questions) than for the teacher-directed instruction (only 261 teacher questions). Activity-based instruction that is characterized by a higher level of teacher-student interactions (teacher questioning or directing students and students responding with some action) is not student-directed activity, but rather it becomes teacher-directed instruction. One cannot rule out that teacher-direction in the "activity-based" treatment of the Bay et al. study was a key factor in achieving the superior results for that treatment. And one certainly cannot use the results of the Bay et al. study (1992) as evidence to support student-directed activity-based instruction over teacher-directed activity-based instruction.

Reference material:

Catrambone, R. & Holyoak, K. (1989). Overcoming contextual limitations on problem-solving behavior. Journal of Experimental Psychology Learning, Memory, and Cognition, 15 (6), 1147-1156.

In a series of six experiments, Catrambone & Holyoak found that a highly directive strategy for problem solving had the greatest level of success. The most effective instruction required students to memorize the ideal answers to 6 questions, one of which was the following verbal description: "If a target is difficult to overcome because a large force cannot be aimed at it from one direction, then divide the force into parts and deliver from many directions." Less explicit descriptions of the deep structural similarities (requiring students only to write answers to the 6 questions and receive corrective feedback) or fewer practice problems were both conditions that resulted in less successful transfer.

Significance of Differences Between Treatment Group Means of Numbers of Principles Known on Third Posttraining Test

Group Pair	Difference Between Means	Standard Error of Difference	t	Level of Significance
Minimum and Intermediate	2.857	.446	6.41	.001
Minimum and Maximum	.941	.4525	2.08	.05
Maximum and Intermediate	1.916	.468	4.09	.001

Treatment-Group Means and Variances for Numbers of Principles Known on Third Posttraining Test

Statistic	Treatment Groups: Minimum	Intermediate	Maximum
Mean	1.733	4.590	2.647
Variances	4.20	4.67	4.93

Catrambone and Holyoak describe the more effective instruction, which requires students to memorize model answers, as "guided discovery."

> Activity 2: Practice reviewing the Tarver and Jung research study found in this issue.

Reviewing a Complex Research Study

Reference material:
Kittell, J. E. (1957). An experimental study of the effect of external direction during learning on transfer and retention of principles. Journal of Educational Psychology, 48, 391-405.

Step 1. What is the research question? Kittel's question is, "What is the effect of differing levels of discovery and explicit teaching on students' future ability to discover?"

Activity-based instruction that is characterized by a higher level of teacher-student interactions (teacher questioning or directing students and students responding with some action) is not student-directed activity, but rather it becomes teacher-directed instruction.

Step 2. Review the test tasks. Three tests evaluate performance on three kinds of tasks: mastery, transfer, and discovery. Reliability is reported as .74, which is within the acceptable range of above .70. (See definition of reliability in glossary.) The reading difficulty of the items was simplified so that the items were unlikely to be tests of reading ability. Each of the three tests evaluated knowledge of 15 principles with three items used to assess each principle (i.e., each test had 45 items). An error on any one of the three items counted as an error for the whole group of three. The principles on the mastery and transfer tests were the instructed principles. The items on the mastery test were the same type as those used in instruction. The items on the transfer test were of a different type but required knowledge of the same 15 principles. The principles on the discovery test were new. The ability to discover new principles is a desirable learning goal.

Step 3. Review the data. Student performance data are reported for three different instructional groups: a "minimum" group that was told only to figure out the principles, an "intermediate" group that was given a verbal statement of the principle, or rule, and a "maximum" group that was told both the principle and the answers during instruction. All students in all groups received immediate feedback as to the correctness or incorrectness of their responses. Pretests were administered which indicated that the three different groups were equivalent before instruction.

A statistical test of significance indicates the level of probability that the differences in means could be explained by chance.

In comparing mean scores, the big question is whether differences are big enough that one can safely conclude the difference was a result of differential treatment, and not a result of random variables. Statistics are used to determine the probability that chance explains the differences in means; therefore, statistical significance is the most important statistic to check for. Kittel's tables reprinted below show three statistics: the mean scores, variances, and statistical analyses for significance for the discovery test (the third posttraining test). (The data for the other tests are not included here.) A standard convention in scientific journals is to portray these three statistics in table form.

Statistical tests of significance generally use the number of subjects in each group, the mean score, and a measure of the spread of scores within the group (either the standard deviation or the variance) in the calculation. The theory behind a test of significance is that random variables affect the performance of students in random ways, thus increasing the performance spread among a group of students. So as the spread in performance within groups increases, the difference in the means must be greater to rule out these random variables as explanations for the difference. If the spread in performance of two groups is very small, then only a small difference in the means may be significant. The measure of spread reported in this study is the variance. The spread is also often reported as standard deviation, which is the square root of the variance.

Statistical tests of significance also correct for the size of the sample. Random variables have less opportunity to be detected in a very small sample, such as 15 in a group. Therefore, a small group must achieve a much greater difference in means to reach statistical significance, whereas a large sample of hundreds of students in a group may achieve statistical significance when there is really very little difference in the performance of the groups. In fact, with very large samples, results may possibly be statistically significant, but not "educationally" significant. That is, the difference may be so small that the small educational gain does not warrant an expensive change in school practices to adopt the more effective treatment. If the results of a statistical test of significance are provided, the reader need not be too concerned about the size of the sample.

A statistical test of significance indicates the level of probability that the differences in means could be explained by chance. The above table from the Kittel study shows that the level of significance in the differences in the means of the intermediate and minimum group was $p < .001$. This means that the probability that chance or other random variables explain the differences in the means is less than 1 in 1000. The cutoff score for an acceptable level of significance is $p < .05$, which means that the probability that random variables explain the differences in the means is less than 5 in 100.

As Kittel should, he interpreted only significant differences in his conclusions: "The Intermediate treatment group was able to discover a significantly greater number of new principles than either of the other two treatment groups. The Maximum treatment group was able to discover a significantly greater number of new principles than the Minimum treatment group" (p. 400). This summary by Kittel seems completely consistent with the information reported in the tables.

The instruction that resulted in better performance on a test of discovery involved telling students a principle or rule and requiring them to apply the rule to examples.

Kittel integrates his findings with those of other studies reviewed in the introduction: "Evidence from this experiment in conjunction with that of similar experiments indicates that furnishing learners with information in the form of underlying principles ... may provide the background enabling future discovery of new principles."

Step 4. Review the procedures. The instruction that resulted in better performance on a test of discovery involved telling students a principle or rule and requiring them to apply the rule to examples. The less effective treatments involved (a) a discovery method, where students were expected to learn to discover by discovering, and (b) a radically explicit method, where students were told both the rule and the answers meaning they did not have to apply the rule.

The procedures in the Kittel study were designed to test a theoretical assumption regarding the relationship between teacher direction and students' ability to discover. For this reason they are not entirely naturalistic but isolate a teaching variable.

Step 5. Compare with your teaching situation. There is no support in this study for using discovery methods to teach students how to become better discoverers. The rule- or principle-teaching procedures used in the Kittel study were simplified to isolate a single theoretical variable. Teachers should feel free to apply the theory (i.e., teach rules) in more fully fleshed out instructional interactions, rather than literally implementing the same basic instructional procedures that Kittel used. The age of the subjects was sixth-grade. A teacher may need to be alert to the fact that the theory may or may not apply to children of other age ranges.

Replications help validate the generalizability and credibility of the findings from one study. In the case of the Kittel study, a number of other studies with a similar design, but with subjects of different ages and abilities, had similar results (Craig, 1956; Wittrock 1963; Wittrock & Twelker, 1964). These replications add credibility to the findings.

References

Bruer, J. (1993). Schools for thought: A science of learning in the classroom. Cambridge, MA: The MIT Press.

Craig, R. C. (1956). Directed versus independent discovery of established relations. *Journal of Educational Psychology, 47*, 223-234.

Prabhu, N. (1990). There is no best method–Why? *TESOL Quarterly, 24*(2), 161-176.

Schoenfeld, A. H. (1987). Cognitive science and mathematics education: An overview. In A. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 1-31). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wittrock, M. C. (1963). Verbal stimuli in concept formation: Learning by discovery. *Journal of Educational Psychology, 54*, 183-190.

Wittrock, M. C., & Twelker, P. A. (1964). Prompting and feedback in the learing, retention, and transfer of concepts. *British Journal of Educational Psychology, 34*, 10-18.

Appendix A

According to scientific research, the following answers are true:

_b_ 1. Are more and better quality ideas produced if (a) students work in groups to brainstorm. (b) students work individually to brainstorm.

Dunette, M., Campbell, J. & Jastaad, K. (1963). The effects of group participation on brainstorming effectiveness for two industrial samples. *Journal of Applied Psychology, 47*, 10-37.

Bouchard, T., & Hare, M. (1970). Size, performance, and potential in brainstorming groups. *Journal of Applied Psychology, 54*, 51-55.

Dillon, P., Graham, W., & Aidells, A. (1972). Brainstorming on a "hot" problem: Effects of training and practice on individual and group performance. *Journal of Applied Psychology, 56*, 487-490.

Bouchard, T. (1971). Whatever happened to brainstorming? *Journal of Creative Behavior, 5*, 182-189.

_b_ 2. Is it more efficient to (a) have students brainstorm first, then develop criteria for evaluating a solution.
(b) have students develop the criteria first, then brainstorm.
(Note: In terms of quantity and quality of ideas, there is no difference between a and b.)

Weisskopf-Joelson, E. & Eliseo, T. (1961). An experimental study of the effectiveness of brainstorming. *Journal of Applied Psychology, 45*, 45-49.

Brilhart, J., & Jochem, L. (1964). Effects of different patterns or outcomes of problem-solving discussions. *Journal of Applied Psychology, 48*, 175-9.

Johnson, D., Parrott, G., & Stratton, R. (1968). Production and judgment of solutions to five problems. *Journal of Educational Psychology Monograph Supplement, 59*, No. 6, Part 2.

Gerlach, V., Schutz, R., Baker, R., & Maser, G. (1964). Effects of variations in test directions on originality in test response. *Journal of Educational Psychology, 55*, 79-83.

See Weisberg, R. (1986). *Creativity: Genius and other myths* for a complete summary of the findings of these studies and a discussion of the marketing efforts that led to the widespread acceptance of beliefs that are not supported by research.

Appendix B

The following table summarizes the steps to reviewing research outlined in this article.

Summary

Sorting rule for finding research relevant to the selection of teaching procedures/tools: In any search of research literature, such as through ERIC, important descriptors to include are: **Research** and **(intervention or experiment)**.

Steps in Reviewing Research

Step 1. What is the research question?
- Identify the independent and dependent variables in the question?

Step 2. Review the test tasks
- Does the task represent a desirable learning goal?
- Were the tests fair and reliable (a reliability coefficient of at least .70)?

Step 3. Review the data
- Were differences in the performance of the groups on the posttests statistically significant at $p < .05$?
- Were the different instructional groups randomly assigned to treatments and equivalent before treatment?
- Is the experimenters' summary of the data consistent with the actual data?

Step 4. Review the procedures

- What procedures / tools were used to achieve the better results?
- How did the procedures used with the more successful group differ from those used with the less successful group?
- What alternative procedures / tools were not compared which might be just as effective, or more effective?
- Do the actual procedures match their descriptive label as you understand it?

Step 5. Compare with your teaching situation

- What age and background were the students in the study? Is there reason to believe the same results would occur in your classroom?
- Was the performance of the better group sufficiently high to warrant your adoption of the better procedure / tool? (Was it higher than what your students already achieve?)
- Can you replicate the successful procedures in your classroom?

Footnote

[1] The comparison with the control group was probably not fair. The measures were clearly designed to align with the instruction that the experimental groups received. Vance and Kieren appropriately do not make much of the poorer performance of the control group.

Glossary

Qualitative research—Research that emphasizes elaborate description of social or instructional settings. Exclusively qualitative research methods are most appropriate in studies where no generalizations or evaluations are desired, for example, a study of the history of Smithville School. Qualitative studies do not provide evidence for generalizations (i.e., theories), though they are sometimes used in education for developing hypotheses. These hypotheses (sometimes called "theories" in education) should be subsequently tested using quantitative methods.

Quantitative research—Research in which numeric data are collected and statistically analyzed to compare across several observations and determine similarities and differences. Quantification is necessary for making comparisons and evaluations. Even descriptors such as *more, less* or *better, worse* are quantitative, as is the identification of the presence (+) or absence (-) of a characteristic. Quantitative research can be both descriptive and experimental. Experimental research, where a teaching variable is manipulated in the study, is the best design for evaluating the relative effectiveness of different teaching tools and practices on learning. Educational studies are generally most informative when both quantitative and qualitative methods are used to describe the results.

Reliability—The degree of consistency in a test over time (test-retest), across raters (interrater), across items within a test (internal consistency), and so on. The type of reliability desired depends on the nature of the construct being measured. To measure a theoretically stable construct, such as intelligence, test-retest reliability is important as a measure of stability. In evaluations involving more subjective impressions, such as performance-based assessment or essay evaluation, interrater reliability is important as a measure of consistency over raters. In measuring performance on a particular skill, such as fractions computation skills or science problem solving, internal consistency reliability is calculated to determine if the test items are sufficient in quantity and quality to measure a construct. Some measures may require more than one type of reliability.

Statistical significance—A condition in which two or more statistics (e.g., means) are more different than would be expected by random variation (chance). This condition is represented as the probability that the difference would have occurred by chance. As a matter of convention, the results of a study can be interpreted using general statements (conclusions) only if the probability that these results occurred by chance is less than five in one hundred ($p < .05$). In other words, if the experiment finding method A worked better than method B were repeated 100 times, method A would probably be better than method B 95 times out of 100. This may seem extremely rigorous, since most teachers would be happy using a method that worked better only 60 or 80 times out of 100. However, the rigorous convention that p must be less than .05 for statistical significance helps compensate for the fact that real experiments can hardly control all variables perfectly.

A Comparison of Mathematics Achievement and Mathematics Attitides of First and Second Graders Instructed With Either a Discovery-Learning Mathematics Curriculum or a Direct Instruction Curriculum

RESEARCH

Sara G. Tarver and Jane S. Jung
Department of Rehabilitation Psychology and Special Education
University of Wisconsin-Madison

Abstract: *This study was conducted to compare the effects of a Direct Instruction mathematics curriculum,* Connecting Mathematics Concepts *(CMC), and a discovery learning mathematics curriculum,* Math Their Way *combined with Cognitively Guided Instruction (MTW/CGI), with students in first and second grades. At the end of second grade, CMC students had significantly higher scores than MTW/CGI students on both the computation and the concepts/applications components of mathematics achievement. In addition, the CMC students had significantly higher scores on a survey of student attitudes towards mathematics. Comparisons of grade equivalents suggest that the Direct Instruction CMC curriculum benefitted high-performing as well as low-performing students.*

"By the year 2000, U.S. students will be first in the world in science and mathematics achievement." That is goal #4 of the National Education Goals contained in Goals 2000 legislation (National Education Goals Panel, 1992).

International comparisons indicate that we have a long way to go to accomplish that lofty goal. Several evaluations of mathematics performance conducted within the last decade revealed substantially lower performance for American students than for students from a number of foreign countries (International Association for the Evaluation of Educational Assessment, 1987; Lapointe, Mead, & Phillips, 1989; McKnight et al., 1987; National Assessment of Educational progress, 1992). Lapointe and colleagues (Lapointe, et al., 1989; Anrig & LaPointe, 1989) reported that American 13-year-olds placed last in mathematics when compared to students in four other countries and four Canadian provinces and that only 16% of American eighth graders have mastered the content of a typical eighth-grade mathematics textbook. Comparisons with Japanese and Chinese students revealed shockingly low mathematics performance for American students (Stevenson, 1990, 1992; Stevenson, Chuansheng, & Lee, 1993).

Mathematics educators have expressed doubt that Americans can reach the goal of being first in mathematics by the year 2000, or even shortly thereafter (e.g., Romberg, 1994). Some contend that the kinds of massive changes that will be required to produce such a monumental turnaround will take at least a generation. Most agree that major curriculum changes will be necessary if we are to meet the dual challenges of (a) teaching higher order mathematical concepts and strategies (e.g., mathematical reasoning and the application of problem solving strategies to real-world problems) as well as basic computation skills and strategies and (b) enabling an increasingly diverse student body to acquire those higher order concepts and skills (Carnine, Jones, & Dixon, 1994).

Although most mathematics educators agree that curriculum reform is essential to the desired turnaround, they are far from agreement regarding the kinds of changes to be made. Age-old debate about the relative merits and demerits of discovery-oriented instructional approaches vs. explicit instructional approaches continues. Although current versions of that debate are strikingly similar to earlier versions (Rappaport, 1976), the debate has intensified since publication of the Curriculum Standards of the National Council of Teachers of Mathematics (NCTM) in 1989.

The 1989 NCTM Standards state that all students should learn to value mathematics, become confident in their ability to do mathematics, become mathematical problem-solvers, learn to communicate mathematically, and learn to reason mathematically. It is widely known that the NCTM Standards promote the constructivist philosophy that is *de rigueur* in educational circles today. It is not so widely known that an accompanying document, Professional Standards for Teaching Mathematics (1991), promotes teaching practices consistent with that philosophy. Together, the 1989 and 1991 documents contain statements of *what* students should learn (i.e.,content) AND examples of *how* students should be taught that content (i.e., teaching methods). The *what* plus the *how* constitutes the instruction that students are to receive. A third Standards document, Assessment Standards for School Mathematics, is to be released in 1995. That third document will contain assessment activities and the criteria by which assessment activities are to be judged. In essence, the tripartite Standards package will define the mathematics content to be taught, the teaching practices to be used to teach that content, and the procedures to be used to assess the effectiveness of the instruction.

Teaching practices included in the 1991 Standards are very similar to practices associated with the Piagetian Cognitive Curriculum (Weikart, Hohmann, & Rhine, 1981) and "new math" (Macarow, 1970; Rappaport, 1976) in the 1960s and 1970s. Critics of the NCTM Standards have raised concerns about the lack of research support for those teaching practices (Carnine et al., 1994; Hofmeister, 1993) Three years after widespread adoption of the teaching standards, a parent requested "data which is not anecdotal, prospective or extrapolative that justifies the adoption of the NCTM practices and methods" In a response to that parent, an NCTM official stated that "data of the type you requested are not yet available" (see Fall, 1994, issue of *Effective School Practices*). NCTM procedures for assessing the effects of the NCTM teaching practices are to be put into place with publication of the Assessment Standards in 1995, four years after widespread adoption of those practices.

A more responsible approach to the identification and promotion of effective mathematics instruction is warranted. Effective instruction should be identified by experimental research that compares viable alternatives on a small scale. Recommendations for widespread adoption should be consistent with findings from small scale experiments. Research that compares discovery-learning approaches with Direct Instruction approaches is especially needed at this point in time to enlighten the debate about reform of the mathematics curriculum.

In this study, a popular discovery-learning curriculum (*Math Their Way* combined with Cognitively Guided Instruction) was compared to a Direct Instruction curriculum (*Connecting Mathematics Concepts*). Effects on mathematics achievement and student attitudes towards mathematics were assessed for first and second graders in one elementary school. The *Math Their Way*/Cognitively Guided Instruction (MTW/CGI) curriculum emphasizes discovery learning through manipulative activities and guidance based on each student's cognitions. The *Connecting Mathematics Concepts* (CMC) curriculum emphasizes explicit explanations of mathematical concepts and connections among the concepts. Although both emphasize application to real-world problems, they do so in very different ways.

In the CMC program, independent application activities follow teacher-directed explanations of concepts and strategies. Basic understandings acquired in teacher-directed instruction are strengthened through application activities which require the student to generalize that which was learned in the preceeding teacher-directed instruction. In MTW/CGI, instruction begins with opportunitites for the student to apply already-existing knowledge and strategies. Student responses to these opportunities provide the basis for the teacher's guidance toward more complex strategies and understandings.

In both approaches, application activities involve the use of manipulatives. In CMC lessons, manipulative activities supplement and enrich teacher-directed instruction and thus come toward the end of the lesson. In MTW/CGI, instruction is based almost solely on manipulative activities. The MTW/CGI teacher observes the student while engaged in manipulative activities, attempts to determine the cognitive strategies employed by each student, and provides guidance that builds on those strategies.

Both CMC and MTW/CGI emphasize mathematical reasoning. In CMC, reasoning strategies are taught directly; in MTW/CGI, opportunities for students to develop reasoning strategies are provided.

Method

Subjects. Subjects were 119 students entering first grade in a midwestern suburban elementary school in September of '92. The 119 students were assigned to one of five classes—one experimental CMC class (N = 26) and four MTW/CGI control classes (Ns = 21,22,25,25). Of those 119 students, 112 were still enrolled in the school and were posttested in May of the first grade year and 88 were still enrolled in the school and were posttested in late March of the second grade year.

Teachers and Teacher Training. One of five first-grade teachers ('92-'93) and one of five second-grade teachers ('93-'94) were assigned to the experimental CMC class. The other four classes combined constituted the control group. Control classes used the mathematics curriculum that was already in place in the school. Teachers described that in-place curriculum as a combination of *Math Their Way*(MTW) and Cognitively Guided Instruction (CGI). All teachers, including the experimental CMC teachers, had both experience using and prior training in the use of the MTW/CGI curriculum. None of the teachers had either experience with or prior training in the use of the CMC program.

Two kinds of training were provided for the two teachers assigned to the experimental CMC class: (a) 4-1/2 hours of training at a summer conference; that training was focused on the curriculum design of CMC and methods of delivering the curriculum, and (b) on-the-job training and consultation by a university supervisor of student teachers.

On-the-job training for the first-grade CMC teacher consisted of the following:

1. Before instruction began in September, the supervisor met with the classroom teacher to coordinate schedules, organize materials, and plan for placement testing.
2. During instruction in September, the supervisor was present for math class each day. Initially, the supervisor demonstrated by teaching entire lessons while the classroom teacher observed. Gradually, the classroom teacher assumed more of the teaching responsibilities, such that by the end of September she had assumed complete responsibility for entire lessons. Throughout the month, the supervisor provided specific feedback on techniques of delivering instructions and monitoring student progress.
3. In October, supervisory visits were reduced to 2-3 per week, during which time the supervisor continued to provide feedback and consultation.
4. For the remainder of the school year, the supervisor visited the classroom approximately once a week to provide support, consultation, and occasional demonstrations.

On-the-job training for the second-grade CMC teacher was less intensive, partially because the students were already accustomed to the CMC program and knew what was to be expected of them. During the first three days of instruction, the supervisor modelled and the teacher observed; thereafter, the teacher assumed full responsibility for lessons. The supervisor continued to visit twice a week in September and once a week for the remainder of the school year to provide feedback and consultation.

Instructional Grouping Practices in the CMC Classroom

In September through December of the first-grade year, teacher-directed instruction was delivered to the class as a whole. Student progress was monitored closely and individual feedback with reteaching was provided to individuals as needed; however, students were not organized into different instructional groups.

In January of the first-grade year, instructional grouping was initiated in an attempt to better meet the needs of a few low-performing students. Low performers continued to receive the same whole-class, teacher-directed instruction along with the other students. Following the whole-class instruction, the low performers received additional teacher-directed instruction while the other students were engaged in independent activities that required application of the skills they had learned. This supplementary instruction was provided by the classroom teacher in the regular classroom.

Instructional grouping was flexible and the composition of the group receiving additional teacher-directed instruction changed frequently. High- and average-performing students who had been absent for several days would often join this group for a day or two to catch up. Some of the low performers made rapid progress and were returned to the group engaged in independent activities. A few low-performing students continued to receive additional teacher-directed instruction throughout the year.

Second grade grouping practices, like those of first grade, were exceedingly flexible.

The first week of second-grade instruction was devoted to review. During that week, a group of 5-6 high performers was identified and it was decided to establish two instructional groups — a high-performing group and an average-performing group. Teacher-directed instruction was delivered to each of these two groups separately by the second-grade teacher. When one group was engaged in teacher-directed instruction, the other was engaged in independent activities.

Second grade grouping practices, like those of first grade, were exceedingly flexible. Shortly after establishment of a high-performing group, other students made progress and were moved into the high-performing group. To assure successful transition, those students participated in the teacher-directed portion of lessons for both the average-performing and the high-performing groups for a period of time. This seemed to work well; in fact, by the end of second grade, 13 of 20 students were in the high-performing group and the remaining 7 were in the average-performing group. Furthermore, the difference in number of lessons completed was not as great as might be expected. At the end of second grade, the high-performing group had completed Lesson #35 of Level C while the average-performing group had completed Lesson #10 of Level C. As will be indicated by the results to be discussed later, the absence of a group labelled "low-performers" at the end of second grade does not signify a misnomer. At the end of second grade, only one student had a Total Mathematics grade equivalent that was below grade level and that one was only four months below grade level.

Mathematics Achievement Measures. The Comprehensive Test of Basic Skills—Mathematics (CTBS-M) was administered on three occasions—as a pretest in September of the first grade year (Level 10, Form A), as a first grade posttest in May of the first grade year (Level 11, Form A), and as a second grade posttest at the end of March of the second grade year (Level 12, Form A). Levels 11 and 12 of the CTBS-M are composed of two subtests—a Concepts and Applications subtest and a Computation subtest. Scores on these two subtests are averaged to yield a Total Mathematics score. Level 10 of the CTBS-M (i.e., the pretest) is composed solely of a Concepts and Applications test; beginning first graders are not expected to compute. Scores obtained for each participant were: Pretest Concepts and Applications, Posttest Concepts and Applications (Grade 1), Posttest Computation (Grade 1), Posttest Total Mathematics (Grade 1), Posttest Concepts and Applications (Grade 2), Posttest Computation (Grade 2), and Posttest Total Mathematics (Grade 2). Raw scores for each of the measures were converted to normal curve equivalents (NCEs) for statistical analysis and to grade equivalents for additional descriptive data.

Student Attitudes Toward Mathematics. An experimenter-constructed math attitudes survey was administered in May of the first grade year and in late March of the second grade year. The survey consisted of six items, with each item being a question or a statement followed by a three-point, Likert-type response format. The first five items were: Math is fun, I am good at story problems, I am good at + and -, I talk about math to my family and friends, and I am smart at math. The three response options for these items were the words "always" (accompanied by a smiling face), "sometimes" (accompanied by a straight face), and "never" (accompanied by a frowning face). The sixth item was "How smart will you be in math next year?", followed by the words "very smart" (accompanied by six hearts), "smart" (accompanied by three hearts), and "not so smart" (accompanied by one heart). Attitude surveys were scored as follows: 3 points for each positive response of "always" with smiling face and "very smart" with six hearts, 2 points for each response of "sometimes" with straight face and "smart" with three hearts, and 1 point for each response of "never" with frowning face and "not so smart" with one heart. Total possible score was 18.

It should be noted that items on the math attitudes survey are relevant to the following NCTM standards: students should learn to value mathematics (e.g., "Math is fun"), students should become confident in their ability to do math (e.g., "I am good at story problems", "I am good at + and -", "I am smart at math"), and students should learn to communicate mathematically ("I talk about math to my family and friends").

Results

Mathematics Achievement

Results of comparisons for end of grade 1 scores and end of grade 2 scores are shown in Tables 1 and 2, respectively. One-way analysis of variance (ANOVA) was used to analyze differences between the CMC group and the MTW/CGI group on all of the measures except Concepts and Applications, in which case one-way analysis of covariance (ANCOVA) was used. Although the Concepts and Applications scores of the two groups did not differ significantly on the pretest, a Pearson-r correlation of .55 for pre- and posttest Concepts and Applications scores at the end of first grade indicated that ANCOVA should be used to adjust statisticially for differences in pretest scores. The usual significance level of .05 was adopted for all analyses.

Mean NCEs, standard deviations, F-values, and p-values are shown in Tables 1 and 2. As indicated by end-of-first- grade scores displayed in Table 1, significant differences in favor of the CMC group were found for Computation and Total Mathematics, but not for Concepts and Applications. Although the first grade CMC group had a somewhat lower mean on pretest Concepts and Applications and a somewhat higher mean on posttest Concepts and Applications, adjusted Concepts and Applications means did not differ significantly, $F(1,109) = 2.510$, $p = .1160$. As indicated by end-of-second-grade scores displayed in Table 2, significant

Table 1

Mean Normal Curve Equivalents, (Standard Deviations), F-Values and p-Values for CMC and MTW/CGI Groups (End of First Grade)

	Pretest	Posttest		
	Concepts and Applications	Concepts and Applications	Computation	Total Mathematics
CMC	41.913	66.261	73.217	70.391
	(16.793)	(14.611)	(18.449)	(15.828)
MTW/CGI	47.921	63.888	52.461	58.449
	(17.764)	(21.187)	(23.396)	(22.249)
$F(1,110)$[a]	2.136		15.563	5.843
$F(1,109)$[b]		2.595		
p	.1467 n.s.	.1101 n.s.	.0001	.0173

[a]analysis of variance. [b]analysis of covariance.

Table 2

Mean Normal Curve Equivalents, (Standard Deviations), F-Values and p-Values for CMC and MTW/CGI Groups (End of Second Grade)

	Pretest	Posttest		
	Concepts and Applications	Concepts and Applications	Computation	Total Mathematics
CMC	42.200	76.150	82.200	79.250
	(17.606)	(20.200)	(20.141)	(19.290)
MTW/CGI	48.544	71.029	51.162	61.279
	(17.833)	(15.999)	(17.503)	(18.449)
$F(1,86)$[a]	1.967		45.352	14.368
$F(1,85)$[b]		7.166		
p	.1644 n.s.	.0089	.0001	.0003

[a]analysis of variance. [b]analysis of covariance.

differences in favor of the CMC group were found for Concepts and Applications as well as Computation and Total Mathematics.

Tables 3, 4, and 5 provide additional descriptive comparisons of the CMC and MTW/CGI groups. In Table 3, mean grade equivalents and grade equivalent ranges are given for the two groups. CMC grade equivalents exceeded those of the MTW/CGI group for all posttest measures (5.6 compared to 4.7 on Concepts and Applications, 3.8 compared to 2.7 for Computation, and 4.5 compared to 3.3 on Total Mathematics). All posttest mean grade equivalents for the CMC group were above the 2.7 grade equivalent that would be expected of the average student at the end of March of the second grade year. Particularly noteworthy is the 5.6 mean grade equivalent for Concepts and Applications.

It should be noted that posttest mean grade equivalents for the MTW/CGI group were also at or above grade level. The 4.7 mean grade equivalent for Concepts and Applications was well above the expected 2.7 grade equivalent. Overall, the grade equivalent data shown in Tables 3-5 show that the performance of both the CMC and the MTW/CGI groups improved more than expected based on national norms.

Both types of instruction benefitted low-performing students; however, the CMC curriculum provided the greater benefits.

Table 4 shows number and percent of students scoring below grade level on the CTBS-M at end of second grade. Because second grade posttesting was conducted in March, grade level for posttesting was defined as 2.7. Grade level for pretesting in September of the first grade year was defined as 1.0. Whereas 60% of CMC students scored below grade level on the Concepts and Applications pretest (compared to 43% of MTW/CGI students), only 5% scored below grade level on the Concepts and Applications posttest (compared to 13% for the MTW/CGI group). Only 10% of the CMC students scored below grade level on the Computation posttest (compared to 32% of MTW/CGI students) and only 5% of the CMC students scored below grade level on the Total Mathematics posttest (compared to 21% of MTW/CGI students). Again, it should be noted that percentage of students scoring below grade level decreased from pretest to posttest for both the CMC and the MTW/CGI groups. This finding suggests that both types of instruction benefitted low-performing students; however, the CMC curriculum provided the greater benefits.

The CMC curriculum enabled more students to reach very high levels of performance.

Table 5 shows number and percent of students scoring at the ceiling (maximum score possible) on the pretest and end-of-second-grade posttest. Percents for the two groups were very similar on the pretest (5% for the CMC group and 6% for the MTW/CGI group), but very different on the posttest (20% for CMC compared to 3% for MTW/CGI students on Concepts and Applications posttest; 50% for CMC compared to 3% for MTW/CGI students on Computation posttest, and 20% for CMC compared to 0% for MTW/CGI students on Total Mathematics posttest). The CMC curriculum enabled more students to reach very high levels of performance.

Student Attitudes Towards Mathematics

Scores on the student attitude towards mathematics survey administered at end of first grade and again at end of second grade were analyzed using one-way ANOVAs. Numbers of students included in the attitude analyses were slightly lower than the numbers included in achievement analyses because a few students who took all achievement tests were not present at the times that the attitude survey was administered. Attitude $\underline{N}$ for first grade was 107 (22 in the CMC group and 85 in the MTW/CGI group), compared to an achievement $\underline{N}$ of 112. Attitude $\underline{N}$ for second grade was 86 (20 in the CMC group and 66 in the MTW/CGI group), compared to an achievement $\underline{N}$ of 88. All students included in attitude analyses were also included in achievement analyses.

Students who received Direct Instruction learned to apply problem solving strategies and reason mathematically at a level that surpassed that of the students receiving the discovery learning curriculum.

Attitude means for the CMC and MTW/CGI groups at end of first grade were 15.591 and 14.659, respectively. First grade means did not differ significantly, $\underline{F}(1,105)=2.561$, $\underline{p}=.1125$. Attitude means for the CMC and MTW/CGI groups at end of second grade were 16.050 and 14.864, respectively. Second

Table 3

Mean Grade Equivalents and (Grade Equivalent Ranges) for CMC and MTW/CGI Groups

	Pretest	Posttests		
	Concepts and Applications	Concepts and Applications	Computation	Total
CMC	.9	5.6	3.8	4.5
	(0.0 - 2.1)	(2.0 - 12.9+)	(2.4 - 4.7)	(2.3 - 7.1)
MTW/CGI	1.1	4.7	2.8	3.3
	(0.0 - 2.1)	(2.1 - 12.9+)	(0.0 - 4.7)	(.8 - 6.3)

Table 4

Number and (Percent) of Students Scoring Below Grade Level for CMC and MTW/CGI Groups

	Pretest[a]	Posttest[b]		
	Concepts and Applications	Concepts and Applications	Computation	Total
CMC	12 (60%)	1 (5%)	2 (10%)	1 (5%)
MTW/CGI	29 (43%)	9 (13%)	22 (32%)	14 (21%)

[a]Grade level = 1.0 grade equivalent. [b]Grade level = 2.7 grade equivalent.

Table 5

Number and (Percent) of Students Scoring at the Ceiling for CMC and MTW/CGI Groups

	Pretest	Posttests		
	Concepts and Applications	Concepts and Applications	Computation	Total
CMC	1 (5%)	4 (20%)	10 (50%)	4 (20%)
MTW/CGI	4 (6%)	2 (3%)	2 (3%)	0 (0%)

grade means did differ significantly, $F(1,84)=6.605$, $p = .0119$.

Discussion

This comparative study of a discovery-learning curriculum and a Direct Instruction curriculum showed significantly higher mathematics achievement for students receiving Direct Instruction. Differences in favor of Direct Instruction were found not only for the computation component of mathematics but for the concepts/applications component as well. This indicates that students who received Direct Instruction learned to apply problem solving strategies and reason mathematically at a level that surpassed that of the students receiving the discovery learning curriculum.

Direct Instruction enables students to reach the NCTM Standards which state that students should learn to value mathematics, become confident in their ability to do mathematics, and learn to communicate mathematically.

Students in the Direct Instruction CMC classroom also developed more positive attitudes towards mathematics. The attitudes findings suggest strongly that Direct Instruction enables students to reach the NCTM Standards which state that students should learn to value mathematics, become confident in their ability to do mathematics, and learn to communicate mathematically.

The CMC curriculum benefitted both high-performing and low-performing students and that the benefits to the high performers were at least as great as the benefits to the low performers.

Evidence to counteract the common claim that Direct Instruction is appropriate for low-performing students but not for high-performing students is provided by the descriptive data in Tables 3-5. Pretest vs. posttest comparisons of mean grade equivalents showed a sharp increase in the percentage of CMC students scoring at the ceiling of the achievement test along with a slight decrease in the percentage of MTW/CGI students scoring at the ceiling. Although pre-post changes in percentage of students scoring below grade level also show greater benefits for the low-performers in the CMC group, relative to those in the MTW/CGI group, both curricula produced decreased numbers of students scoring below grade level.

Overall, the descriptive data suggest that the CMC curriculum benefitted both high-performing and low-performing students and that the benefits to the high performers were at least as great as the benefits to the low performers. Clearly, the benefits of the CMC Direct Instruction mathematics program were limited to neither low- nor high-performing students.

Direct Instruction is much more likely to meet the needs of diverse students grouped together in the regular classroom than are constructivist teaching approaches which are being touted as the solution to diverse students' needs.

Results of both the statistical analyses and the descriptive analyses of this study support the contention that Direct Instruction is a viable alternative for achieving the goals stated in the 1989 NCTM Standards. They also suggest that Direct Instruction teaching practices provide a more effective means of reaching those goals than do the discovery-learning practices recommended in the 1991 Standards document. Finally, they suggest that Direct Instruction is much more likely to meet the needs of diverse students grouped together in the regular classroom than are constructivist teaching approaches which are being touted as the solution to diverse students' needs. Results should speak louder than rhetoric.

References

Anrig, G. R., & LaPointe, A. E. (1989). What we know about what students don't know. *Educational Leadership, 47*(3), 4-9.

Carnine, D., Jones, E. D., & Dixon, R. C. (1994). Mathematics: Educational tools for diverse learners. *School Psychology Review, 23*, 406-427.

Hofmeister, A. (1993). Elitism and reform in school mathematics. *Remedial and Special Education, 14*(6), 8-13.

International Association for the Evaluation of Educational Achievement. (1987). *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective.* Urbana, IL: University of Illinois.

LaPointe, A. E., Mead, N. A., & Phillips, G. W. (1989). *A world of difference: An international* *assessment of mathematics and science.* Princeton, NJ: Educational Testing Service.

Macarow, L. (1970). New math. *School Science and Mathematics, 70*(5), 395-397.

McKnight, C., Crosswhite, F., Dossey, J., Kifer, E., Swafford, J., Travers, K., & Cooney, T. (1987). *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective.* Champaign, IL: Stipes Publishing Co.

National Assessment of Educational Progress. (1992). *NEAP 1992 mathematics report card for the nation and the states.* Washington, DC: National Center for Educational Statistics, Report No. 23-ST02.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Rappaport, D. (1976). The new math and its aftermath. *School Science and Mathematics, 76*(7), 563-570.

Romberg, T. A. (1994). A path to change. *Focus in Change, 16*, 7-9. University of Wisconsin-Madison: The National Center for Effective Schools.

Stevenson, H. (1992). Learning from Asian schools. *Scientific American, 267*(6), 70-76.

Stevenson, H. W., Chuansheng, C., & Lee, S. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science, 259*, 53-58.

Weikart, D. P., Hohmann, C. F., & Rhine, W. R. (1981). High/Scope Cognitively Oriented Curriculum Model. In W. R. Rhine (Ed.), *Making Schools More Effective.* New York: Academic Press. Pp. 201-247.

Send your letter of nomination for the Excellence in Education Awards:

- ☆ **Outstanding Administrator**
- ★ **Outstanding Teacher**
- ☆ **Outstanding Researcher**
- ★ **Distinguished Service**

In your letter describe achievements and provide testimonials, anecdotes and any other relevant information. Gather signatures of other staff members to support your letter. Any information documenting effectiveness is relevant.

Send nominations to: ADI
Awards
PO 10252
Eugene, OR 97440

☞ ☞ Calling All Tapes ☜ ☜

The ADI products division is interested in your video tapes. We're not interested in the "America's funniest home video" type, but video tapes that you use as examples of good instruction, kids working hard, and so on.

Please call 1-206-754-1122 if you have something you think we might be interested in.

Join a local ADI chapter

The persons below are organizing local ADI chapters. They plan to form local support groups and to sponsor local workshops, discussion groups, and newsletters. Contact the person nearest you for more information on local chapters. If your name is not on the list and you would like to form a local chapter, contact ADI, PO 10252, Eugene, OR 97440 or call (503) 485-1293.

Carolyn Cittamlet
1422 S. 13th St.
Philadelphi, PA 19147
Fax: 215-551-9790

Susan Kandell
212 S. Woodhams St.
Plainwell, MI 49080-1753

Kathleen Schaefer
2668 Tareyton Cr.
Stoughton, WI 53589

Patti Clark
Phoenix Academy
11032 Oak St.
Omaha, NE 68144

Paul Koeltzow
10318 Fern Dale Rd.
Dallas, TX 75238
214-341-5373

Diana Morgan/Thaddeus Lott
Wesley Elementary
800 Dillard St
Houston, TX 77091

Clark Walker
300 West 100700
Ftu Green UT 84632

Ken Traupman
248 Nutmeg St.
San Diego, CA 92103

Anna Mae Gazo
3027 Ellen Ct.
Marina, CA 93933

Cathy Watkins
1956 La Linda Ct.
Turlock, CA 95380
cwatkins@koko.csustan.edu

Ursula Garrett
PO Box 241, Apt 169
Kahuku, Hawaii 96731

Chuck Main
PO Box 8
Silverdale, WA 98303

Betty Williams
Dept. of Special Education
AD Box 25
Gonzaga University
Spokane, WA 99258

Babette Engel
343 Dungeness Meadows
Sequim, WA 98382

Helen Munson, Tricia Walsh-Caughlan
1603 NW 41st Circle
Camas, WA 98607

Larry Chamberlain
1063 Stelly's X Rd.
Brentwood 1324,
Vosiao, BC

Dorothy Ross
Terry Fox Sr. Secondary
3550 Wellington
Port Coquitlam, B.C. V3B 3Y5

Vicky Vachon
148 Wolfrey Ave.
Toronto, Ontario M4K 1L3

NEW!

WE'RE ON THE INFORMATION SUPERHIGHWAY

To subscribe to *Effective School Practices* electronic discussion group, or "list," send the following message from your e-mail account:

To: Mailserv@oregon.uoregon.edu
Message: Subscribe effschprac
(Don't add *Please* or any other words to your message. It will only cause errors. Mailserv is a computer, not a person. No one reads your subscription request.)

By subscribing to the *EFFSCHPRAC* list, you will be able to participate in e-mail discussions of topics of interest to ADI members. You will automatically receive in your e-mail box all messages that are sent to the list. You can also send your news out to the list subscribers, like this:

To: Effschprac@oregon.uoregon.edu
Subj: *Whatever describes your topic.*
Message: *Whatever you want to say.*

Do you have a special need? Perhaps there is someone in the discussion group who has an answer or can help. Are you looking for a job where you can use your DI skills? Or are you looking for DI teachers to employ? Do you have a news flash? Send it all the *EFFSCHPRAC* list.

CONTRIBUTOR'S GUIDELINES

Effective School Practices provides practitioners and decision-makers with the latest research and development news on effective teaching tools and practices. The journal emphasizes practical knowledge and products that have proven superior through scientific testing. Readers are invited to contribute to several different columns and departments that will appear regularly:

FROM THE FIELD: Submit letters describing your thrills and frustrations, problems and successes, and so on. A number of experts are available who may be able to offer helpful solutions and recommendations to persons seeking advice.

NEWS: Report news of interest to ADI's membership

SUCCESS STORIES: Send your stories about successful instruction. These can be short, anecdotal pieces.

PERSPECTIVE: Submit critiques and perspective essays about a theme of current interest, such as: school restructuring, the ungraded classroom, cooperative learning, site-based management, learning styles, heterogeneous grouping, Regular Ed Initiative and the law, and so on.

RESEARCH STUDIES: Present data from your classroom or the results of scientific research. The data should guide other practitioners and decision-makers in evaluating alternative options for school reform.

TRANSLATING RESEARCH INTO PRACTICE Integrate a larger body of empirical research into a defined practice that can be implemented in schools.

BOOK NOTES: Review a book of interest to members.

NEW PRODUCTS: Descriptions of new products that are available will be featured. Send the description with a sample of the product or a research report validating its effectiveness. Space will be given only to products that have been field-tested and empirically validated.

LIST OF DEMONSTRATION SITES: We wish to maintain an on-going list of school sites with exemplary implementations and impressive student outcomes. Submit the name of the exemplary school or classrooms, the names of the programs being implemented, and contact information so that visitations may be arranged.

TIPS FOR TEACHERS: Practical, short products that a teacher can copy and use immediately. This might be advice for solving a specific but pervasive problem, a data-keeping form, a single format that would successfully teach something meaningful and impress teachers with the effectiveness and cleverness of Direct Instruction.

MANUSCRIPT PREPARATION

Authors should prepare manuscripts according to the fourth revised edition of the *Publication Manual of the American Psychological Association*, published in 1995. Copies may be ordered from:

Order Department
American Psychological Association
1200 Seventh St., N.W.
Washington, DC 20036

Send an electronic copy, if possible, with a hardcopy of the manuscript. Indicate the name of the word-processing program you use. Save drawings and figures in separate files. Electronic copy should replace text that is underlined according to the APA format, with italic text.

Illustrations and Figures: Please send drawings or figures in a camera-ready form, even though you may also include them in electronic form.

Completed manuscripts should be sent to:

Bonnie Grossen, Ph.D.
Editor, *Effective School Practices*
PO Box 10252
Eugene, OR 97440

Acknowledgement of receipt of the manuscript will be sent by mail. Articles are initially screened by the editor for content appropriateness, then sent out for review by peers in the field. These reviewers may recommend acceptance as is, revision without further review, revision with a subsequent review, or rejection. The author is usually notified about the status of the article within a 6- to 8-week period. If the article is published, the author will receive five complimentary copies of the issue in which his or her article appears.

Order Back Issues/Monographs on Important Themes

All back issues and monographs are $5.00. Postage and handling is $3.00 per order of any size. Prepaid orders only.

Discriminatory Educational Practices **$5.00**
Effective School Practices, Spring, 1993, Volume 12, No. 2

ABSTRACT: Research has documented discriminatory effects for two popular school reforms: whole language and "developmentally appropriate practice" as it has been defined by the National Association for the Education of Young Children. This edition summarizes the research evaluating effects of these reforms on the upward mobility and learning of economically disadvantaged children, minority children, and special education children. These diverse learners in programs incorporating the popular "child-centered" pedagogies are less likely to acquire the tools they will need for economic success and have lower self-esteem than children in traditional programs.

Heterogeneous Grouping and Curriculum Design **$5.00**
Effective School Practices, Winter, 1993, Volume 12, No. 1

ABSTRACT: Heterogeneous grouping is a superficial and ineffective solution to the problem of discrimination in education. Equal access to education involves much more than having equal access to a seat in the classroom. This edition presents research summaries and perspectives surrounding grouping decisions. Research finds subject-specific homogeneous grouping most effective in subjects that are skills-based, such as reading and mathematics. The reprinted education survey by the *Economist* compares educational systems around the world and finds America's attempt to provide equal education for all a failed experiment. The *Economist* praises Germany's ability to turn out the most highly skilled workers in the world. Both *Forbes* and the *Economist* criticize many of the currently popular American reforms, such as whole language and heterogeneous grouping, for the mediocrity they seem to encourage.

Listing of Effective Programs **$5.00**
Effective School Practices, monograph, 1993, also *ADI News*, Volume 12, No. 5.

ABSTRACT: This issue features a complete annotated listing of Direct Instruction, programs authored by Zig Engelmann and his colleagues. Also included are procedures for obtaining funding, addresses of funding sources, and a model proposal.

Wholistic Approaches **$5.00**
ADI News, Summer, 1992, Volume 11, No. 4

ABSTRACT: Effective instruction (e.g., Direct Instruction,) provides wholistic integration of skills that have been specifically taught. Wholistic programs that do not teach important component skills are inferior. A study is reported that shows that students learning from Direct Instruction programs in mathematics achieve higher scores than students learning from the new teaching standards promoted by National Council of Teachers of Mathematics. A synthesis of studies in reading shows that using Direct Instruction reading programs result in higher reading scores than whole language programs that provide no instruction in component skills, such as decoding.

ADI News, **Volume 11, No. 2** **$5.00**

ABSTRACT: This edition includes a study comparing the effects of four procedures for parents to use in teaching reading to their children. Parents using *Teach Your Child to Read in 100 Easy Lessons* (see ADI materials list for ordering information) obtained the highest reading improvement scores with their children. This edition also reports a comparison of the achievement scores of Wesley Elementary, a Direct Instruction school, with ten other schools, the results of a comparison of meaning-based versus code-based programs in California, and other reports of the effectiveness of Direct Instruction programs with special populations.

Historical Issue III **$5.00**
ADI News, Volume 8, No. 4

ABSTRACT: The historical series reprint highlight articles and contributions from earlier editions. The featured articles in this edition are divided into the following sections: (1) Implementation strategies and issues, (2) Direct Instruction research studies, and (3) Research related to DI's goals. Russell Gersten's response to a study that is widely discussed among promoters of the current child-directed instruction reform is reprinted in this edition. That study by Schweinhart, Weikart, and Larner is highly critical of DI preschool programs. Gersten criticizes that study primarily for using self-report data to evaluate delinquency and for interpreting nonsignificant differences as if they were significant.

Historical Issue I **$5.00**
ADI News, Volume 7, No. 4.

ABSTRACT: The featured articles in this issue are divided into the following sections: (1) Introduction, (2) Research studies, and (3) Management strategies. These include a classic essay by Zig Engelmann "On Observing Learning," a high school follow-up study on Follow Through children in Uvalde TX, a meta-analysis of the effects of DI in special education by W.A.T. White, and other studies reporting the effects of DI in teaching English as a Second Language, poverty level preschoolers, secondary students, and moderately retarded children. Also included are classroom management tips from Randy Sprick and Geoff Colvin, along with a school-wide discipline plan.

Beginning Reading Instruction $5.00
Effective School Practices, Winter 1994, Volume 13, No. 1

ABSTRACT: Research still shows that systematic phonics instruction with a code-based reader are important components of effective initial reading instruction and are not incompatible with most whole language activities. Read Keith Stanovich's analysis of reading instruction issues in *Romance and reality* and Patrick Groff's review of *Reading Recovery* research. Read how a highly successful school teaches reading to Spanish-speaking children. Edward Fry also provides a set of tools for solving common reading problems.

Achieving Higher Standards in Mathematics $5.00
Effective School Practices, Spring 1994, Volume 13, No. 2

ABSTRACT: The standards from the National Council of Teachers of Mathematics prescribe teaching practice more than they set standards for student performance. Several research articles provide evidence that the NCTM teaching practices are probably not the best practices for achieving the student performance standards implied in the standards.

OBE and World Class Standards $5.00
Effective School Practices, Summer 1994, Volume 13, No. 3

ABSTRACT: This issue is a critique of outcome-based education. Criticisms from educational researchers and from the American Federation of Teachers are featured. Positive suggestions for education reform legislation are offered, as well as some guidelines for evaluating standards. The standards of most states are criticized for their lack of rigor, for their non-academic focus, and for their evaluation systems that do not provide information regarding the effectiveness of the school programs, but rather only evaluate individual students.

Twenty Years of Effective Teaching $5.00
Effective School Practices, Fall 1994, Volume 13, No. 4

ABSTRACT: Two keynote addresses by Sara Tarver and Jean Osborn at the summer conference provide an overview of the history of Direct Instruction. Headline news articles featuring Direct Instruction and/or disappointing results from trendy approaches are reprinted. An exchange of letters between a Montana parent and the National Council of Teachers of Mathematics highlights issues regarding school adoption of unproven, faddish methods, textbooks, and philosophies. The NCTM is unable to provide evidence that the teaching methods they promote improve learning. NCTM claims there are no measures that assess the kinds of outcomes they wish to achieve. They expect to have a guide for assessment published in 1995, 4 years after the guide for teaching practice was published. The Montana parent argues that the assessment should be used to evaluate the practices before they are promoted nationwide.

Add $3.00 postage & handling *per order*.
Prepaid orders only.

$ $ $ $ $ $ $ $ $ $ $ $ $ $

ADVERTISING POLICY

The Association for Direct Instruction advertises only products that are endorsed by ADI. ADI endorses only products that consistently produce significant learning and behavioral outcomes. In order to obtain ADI's endorsement, vendors should send a sample of the product and copies of research articles, comparative studies, pre-posttest data, or any other data verifying the product's effectiveness, for the ADI Board to review. Information about teacher satisfaction is insufficient.

After obtaining ADI's endorsement, vendors may advertise in ADI publications, may purchase ADI mailing lists, and may display at ADI-sponsored or endorsed conferences. Our advertising rates are $200 for a full-page ad, $125 for a half-page ad, and $75 for a quarter-page ad in our quarterly publication *Effective School Practices*. Products that ADI endorses will also be promoted without charge in the Fall listing of *Effective Programs and Products*. New products will be featured in the "New Products" column of our *Effective School Practices* publication, also without charge. We have mailing lists of varying sizes. The charge for our 3000-name list is $75.

Summer Direct Instruction Training Opportunities

July 17–20
11th Atlantic Coast Conference on Effective Teaching and Direct Instruction
Cape Henelopen High School • Lewes, Delaware
Contact: ACCDI, PO Box 997, Rehoboth Beach, DE 19971

July 23–27
21st Annual Eugene Direct Instruction Conference
"World Class Standards for the 21st Century"
Eugene, Oregon
New Feature: Sunday, July 23rd Pre-Conference Sessions:
Classroom Management–Randy Sprick
Study Skills–Anita Archer
Becoming a Direct Instruction Trainer–Team of ADI Lead Trainers
Contact: ADI, PO Box 10252, Eugene, OR 97440

July 31–August 2
Salt Lake DI Conference
Hilton Hotel • Salt Lake City, Utah
Contact: Richard West, SRA, 10924 S. Shelbrooke Dr, South Jordan, UT 84095

August 14–16
1995 Wisconsin Summer Conference on Effective Instruction
University of Wisconsin-Madison • Madison, Wisconsin
Contact: Chris Dzemske, Wisconsin Center, Room 105
702 Langdon Street, Madison, WI 53706

August 16–18
DI Summer Institute
"Achieving a Balance in an Integrated Classroom"
Seattle Pacific University • Seattle, Washington
Contact: Willy Ertsgaard, 2665 NW 95th, Seattle, WA 98117

October 26–27
21st Carmel Direct Instruction Conference
Carmel Mission Inn • Carmel, California
Contact: Wes Robb, 6527 N Colonial Ave, Fresno, CA 93704

Recommended Resources

School's Out: The Catastrophe in Public Education and What We Can Do About It (1993) by Andrew Nikiforuk.
ISBN: 0-921912-48-X
Price: $19.95 from Macfarlane Walter & Ross
37A Hazelton Avenue
Toronto, CA M5R 2E3
Ask for it at your local bookstore.

Beginning to Read: Thinking and Learning About Print (1990) by Marilyn Jager Adams (A summary by the Center on Reading).
Price: $7.50
Mail orders to: Center for the Study of Reading
University of Illinois
51 Gerty Cr.
Champaign, IL 61820

Direct Instruction Reading (Revised, 1990)
by Douglas Carnine, Jerry Silbert, & Ed Kameenui.
Price: $40.00
Order from: MacMillan Publishing
1-800-257-5755
ISBN: 0-675-21014-3

Antisocial Behavior in Schools: Strategies and Best Practices (1995) by Hill Walker, Geoff Colvin, & Elizabeth Ramsey.
Price: $28.70
Order from: Brooks/Cole Publishing Co.
1-408-373-0728 (ext 137)
Fax: 1-408-375-6414
Email:
adrienne_carter@brookscole.com
(Complimentary copies sent for review for college course. Send request on letterhead.)

Failing Grades (Video) and Annotated Bibliography (1993) featuring Joe Freedman, M.D. & Mark Holmes, Ph.D.
Price: $17.95
Order from: Society for Advancing Research
c/o VICOM Limited
11603-165 Street
Edmonton, Alberta
CANADA T5M 3Z1

If Learning Is So Natural, Why Am I Going To School? (1994) by Andrew Nikiforuk.
Price: $16.99 from Penguin
ISBN: 0-14-02.4264-3
Ask for it at your local bookstore.

Becoming a Nation of Readers (1985)
The Report of the Commission on Reading.
Price: $7.50
Mail orders to: Center for the Study of Reading
University of Illinois
51 Gerty Cr.
Champaign, IL 61820

Direct Instruction Mathematics (Revised, 1990) by Jerry Silbert, Douglas Carnine, & Marcy Stein.
Price: $40.00
Order from: MacMillan Publishing
1-800-257-5755
ISBN: 0-675-21208-1

Interventions for Achievement and Behavior Problems (1991) by 74 contributors, edited by Gary Stoner, Mark Shinn, & Hill Walker.
Price: $52.00
Order from:
National Association of School Psychologists
8455 Colesville Road, Suite 1000
Silver Spring, MD
ISBN: 0-932955-15-0

Effective School Practices
ADI
P.O. Box 10252
Eugene, OR 97440

Non-Profit
Organization
U.S. Postage
PAID
Permit No. 122
Eugene, OR

Dear friend:

This is a complimentary copy of our quarterly publication *EFFECTIVE SCHOOL PRACTICES.* I hope you find the contents helpful and informative. If you have not already joined the *Association for Direct Instruction,* consider joining today.

The mission of the *Association for Direct Instruction* is to improve education. Currently, a major obstacle to improved education is old fads that are recycled as "new reforms," in spite of their failure in the past. Future ADI publications will define instruction that represents the state-of-the-art in research-based practices—instruction that emphasizes thoughtful balance rather than faddish extremes.

Join us! Help counteract faddism in education.
Our membership fee is only $20.00.

With your membership you will receive a subscription to *EFFECTIVE SCHOOL PRACTICES,* and discounts on publications and on staff development workshops. We are also planning a series of monographs on specialized topics. ADI members will receive free copies of these special ADI monographs at the time of printing.

Bonnie Grossen,
Editor, *Effective School Practices*