

Reforming the Wannabe Reformers

Why Education Reforms Almost Always End Up Making Things Worse

Mr. Pogrow uses his own experience and the history of education reform over the past 100 years to argue that the biggest problem in education is with the reformers themselves and with the academicians and researchers who develop the ideas and rationales for the reformers' pet reforms.

BY STANLEY POGROW

FOR MORE THAN a decade, we have been buried in proposed reforms. Those responsible for this avalanche of reforms have taken the perspective that there are problems with the education establishment, problems with society, problems with the political structure, problems with current practice — in short, problems with everything except reformers and their proposed reforms.

Reformers typically feel that their solutions would work if only people would get on board. When people do not jump on board, the reformers conclude that the practitioners are at fault or that the society is at fault for socializing individuals in ways that prevent them from appreciating the wisdom of the reforms. When the proposed reforms, lacking popular support, inevitably end up not working, the refrain is that the reforms were implemented in style but not in substance. Once again, the practitioners are at fault.

I will argue that the biggest problem in education is with the reformers themselves and with the academicians and researchers who develop the ideas and rationales for the reformers' pet reforms. My conclusions stem from my own experience as a reformer, an academician, and the developer of the HOTS (Higher Order Thinking Skills) program¹ and from the history of education reform over the past 100 years.

The State of Education Reform

This appears to be a time when reform is blossoming. The *Kappan* and other major education publications have highlighted dozens of reforms. Examples of current widely advocated reforms include whole language, vouchers, heterogeneous grouping, teacher empowerment, authentic assessment, and team teaching. Lovers of reform are ecstatic; the traditionalists seem to be on the run.

But this isn't the first instance of hyper-reform in the history of education. Another such period ran from the mid-1960s to the mid-1970s. Almost none of the widely advocated reforms of that period — open space, individualization, community-based education — survived.

Unfortunately, the fate of these earlier reforms is typical. The history of education reform is one of consistent failure of major reforms to survive and become institutionalized. David Tyack and his colleagues have found that education reforms typically do not last very long. Larry Cuban has noted that the few that do survive are shorn of their ambitious goals and ideals, becoming instead routinized incremental changes to what exists. The only innovations that survive are those that are high-

ly structural in nature, that are easily monitored (e.g., the Carnegie unit), or that create new constituencies. Cuban refers to the historical success of attempted curriculum reform as "pitiful."² Indeed, reports of research on the innovations of the late 1980s and early 1990s are starting to appear and are generally disclosing failure.³

Does the consistent failure of all but the simplest reforms suggest that educators are stupid, lazy, unimaginative, and uncaring? No! The record of innovation in education is the same as that in other areas. In *Innovation and Entrepreneurship*, Peter Drucker draws on a wide range of human experience to determine the fundamental conditions under which new ideas become successful and enduring innovations in any field. He finds that, historically, the vast majority of innovative ideas and changes throughout human experience have failed to take root. Most remain just interesting ideas. Drucker arrives at the following conclusions regarding the fate of new ideas:

- ideas that become successful innovations represent a solution that is clearly definable, is simple, and includes a complete system for implementation and dissemination;

- successful innovations start small and try to do one specific thing; and
- knowledge-based innovations are least likely to succeed and can succeed only if *all* the needed knowledge is available.⁴

Drucker's conditions for success are a chronicle of human tolerance for uncertainty and ambiguity. Unfortunately, these conditions are violated by virtually every new idea for change that is currently sweeping through the education profession. For example, school restructuring violates all of Drucker's principles. Process learning approaches, in which teachers are left to invent their own methods, violate the first principle. Indeed, students in my educational administration classes decided that only one current reform met all three of Drucker's conditions — integrated social service centers in schools.

Educational practitioners are no less skilled in implementing innovations than practitioners in other professions. The fault lies in the types of reforms they are seduced into pursuing by a reform/academ-

STANLEY POGROW is an associate professor of educational administration at the University of Arizona, Tucson, where he specializes in instructional and administrative uses of computers. He is the developer of the HOTS and SUPERMATH programs.

ic/research community that is largely out of touch with reality.

Myths of the Reform Process

Education reforms almost always fail because they are usually based on combinations of a number of myths.

Myth 1. You can change instruction via advocacy, inservice, and training. The single biggest tool in promoting reform has been advocacy—followed up by massive doses of conferences, inservice training, and university courses. The scenario goes like this: a sense of urgency is created, and a new terminology is coined; a national fellowship develops among the believers; stories of success appear in a journal such as this; and a massive national network of training is created. The advocacy is driven largely by philosophy, with only a smidgen of technique or research supporting the idea. The word then goes out that the technique is supported by research. In retrospect, the supposed research is never very convincing, and the reforms fade away for lack of a real methodology for implementing them. Some examples of current reforms that are built primarily around advocacy and training are the middle school movement, schoolwide approaches to Title I, heterogeneous grouping, full inclusion, and a thinking approach to mathematics.

Reality 1. Large-scale reform requires highly specific, systematic, and structural methodologies with supporting materials of tremendously high quality. (Such methodologies are hereafter referred to as "technologies.") All the inservice training, editorials, and articles cannot make up for the absence of a powerful, yet simple, supporting technology. (Technology is a systematic way of doing something consistently and can be either a specific social process or some specific equipment.) Without such technology, almost all training is a waste of time.

For example, consider the case of middle school reform. Thousands of articles and speeches have advocated the development of child-centered curricula. I recently completed a three-year study to identify exemplary middle school curricula. There were few examples of exemplary curricula to be found.⁵ The only exemplary math curriculum was 20 years old, and the only exemplary comprehensive science curriculum was Canadian. There

were no exemplary comprehensive language arts or social studies curricula. While everyone has been philosophizing about what middle school curricula should look like, no one has bothered to develop them, despite 40 years of advocacy. The basic tools that are needed aren't there; there is a lot of bull but no beef.

The middle school movement is not a singular example. The reality for teachers and principals is that exemplary programs usually do not exist for the goals schoolpeople are being asked to achieve.

Myth 2. Theory is a useful guide for the design of programs and reforms. This is the most cherished belief of the academic and reform community.

Reality 2. Metaphor is much more important to the design of sophisticated programs than research and theory. The key to developing successful programs is to have the right metaphor. The key metaphor in the HOTS program is the "dinner table conversation in the home." This metaphor was the basis for at least 80% of the design decisions. For example, since dinner table conversation is not linked to formal content, it was decided that HOTS would not link to the regular classroom curriculum—a counterintuitive decision for a Title I program. Once you have the right metaphor, though, theory can fill in some of the gaps and help with perhaps 10% to 15% of the decisions that have to be made.

Myth 3. You can reform education by disseminating knowledge and leaving it up to practitioners to apply that knowledge. The REsearch/Academic/Reform community mentioned above (hereafter referred to as REAR) continually claims that it knows what works, if only others would apply that knowledge. For example, Carl Kaestle found that key researchers blamed "the awful reputation of education research" on how the research is disseminated and on the lack of incentives for practitioners to use research.⁶

The knowledge disseminated by REAR consists primarily of advocacy and general theory. The feeling is widespread in the REAR community that its responsibility is to produce general theory and that it is up to practitioners to figure out how to apply that theory. For example, in describing the success of teachers in designing *libritos* based on knowledge of literacy to develop the reading skills of minority students, Claude Goldenberg and

Ronald Gallimore view the local knowledge needed to make implementation decisions as separate from research knowledge.⁷ Essentially, they argue that local practitioners must apply theory to develop their own interventions because knowledge of details will invariably be different for each local site.

Reality 3. Reform requires technology, methodology, structure, dosages, and materials. It is far more difficult to figure out how to implement theory than it is to generate it. I am reasonably intelligent, and it took me 14 years of almost full-time effort to figure out how to consistently work just four thinking skills into a detailed and effective curriculum. Thus it makes no sense to expect practitioners to develop their own techniques for implementing a complex reform idea. While there are many talented teachers who can come up with highly innovative techniques, it's too demanding and too much a hit-or-miss proposition. This is not a criticism of educators. No other field expects its practitioners to develop the techniques that they practice. Indeed, in medicine, if individual practitioners invent their own procedures, we call it "malpractice."

The equivalent of expecting teachers to develop the interventions they are going to apply would be asking an actor to perform Shakespeare—but to write the play first. The role of actors is to make the playwright's lines come alive, not to write those lines. The primary role of teachers is to teach, not to develop their own interventions because the REAR community prefers to philosophize and preach. Professional behavior is judged by the quality with which practitioners implement established procedures, not by whether they can invent them.

The simple fact of the matter is that what Goldenberg and Gallimore called local knowledge is as important in education as in medicine. We do need to know how much and what types of services specific students need to improve their performance; we don't need general philosophical statements such as "All students can learn." My own experience is that it is indeed possible for the right type of research to develop techniques and determine implementation details that are applicable to most local conditions—if REAR is so disposed.

But the bottom line is that no amount of advocacy will cause an innovation to

succeed if it lacks an underlying technology. The individualized education movement of the 1970s is a classic example of a reform with absolutely no technology that most teachers found simply impossible to carry out. We are repeating this mistake today. Current REAR reforms that have little or no underlying technology include restructuring, site-based management, full inclusion, constructivism, and portfolio assessment.

Myth 4. The most important change involves radical reformulation of existing practice, i.e., new paradigms. Whenever a new reform idea is presented, it is usually made to sound revolutionary. The most popular phrase used to describe these new ideas is "paradigm shift." The concept of a paradigm shift can be traced to the work of Thomas Kuhn, who has documented how important periodic reformulations are to the evolution of scientific knowledge.⁸ In education, most reforms are presented as paradigm shifts. Authentic testing is presented as a paradigm shift away from the evil standardized tests; whole language is a paradigm shift away from the evils of phonics; of course, the biggest paradigm shift of all is restructuring itself (whatever that is).

Reality 4. The most important changes are incremental ones. While paradigm shifts are important in the evolution of knowledge, they are extremely rare. Most fields do not have even one per century. Moreover, they are seldom involved in the creation of breakthrough products. Indeed, most lucrative patents and products are incremental refinements of existing technologies. HOTS did not come from any new paradigm; it came from more than a dozen years of tinkering with combinations of new pieces of technology and 2,000-year-old ideas, as well as lots of observation of teachers and students conversing.

Myth 5. The best way to achieve reform is through schoolwide change/restructuring. The representatives of REAR always start with the assumption that it is critical to change whole schools and systems.

Reality 5. Schoolwide change, while a nice idea, has never worked on a large scale and is probably not necessary. In some cases entire schools may need to change, but there is no evidence that entire schools can be changed on a large scale. In a sense, schoolwide change has become a convenient rallying cry that often provides a smokescreen to cover the failure to deal

substantively with the real issues.

Nor is there evidence that whole schools generally need to be changed. My own work suggests that it is possible to produce high levels of learning by providing exemplary activities for just a small part of the day. If we cannot figure out how to provide exemplary learning experiences to every student for an hour a day, we certainly cannot figure out how to consistently change whole schools in substantive ways.

Myth 6. You can develop learning through reforms designed to enhance correlates of learning, such as self-concept or empowerment. We keep creating reforms that focus on everything but learning. For example, in the 1980s we became absorbed in developing student self-concept, computer literacy, and computer equity. So far in the 1990s we are absorbed with detracking, empowering, eliminating labels, sex equity, changing tests, and increasing democratic participation. While all of these are important and are related to learning, they are not learning.

Reality 6. The best way to enhance learning is to develop more powerful programs to enhance learning. Movements built around the correlates of learning never lead to substantial improvements in learning. These correlates should be produced as by-products of increased learning. Unfortunately, we always seem to get sidetracked and end up trying to produce the correlates directly. It is amazing how many different reforms we have distracted ourselves with over the years. (I strongly suspect that teacher empowerment and authentic testing are similar diversions.) REAR is very inventive about developing new ways to engage us in pursuit of side-show issues. Indeed, REAR often seems to have little interest in developing or defining systematic learning environments.¹⁰

Myth 7. You can understand large-scale change by understanding what happens on a very small scale. This is the biggest myth of all! Physicists understand that physical processes at the small-scale, individual subatomic particle level are very different from those at the large-scale, aggregate human level. The fundamental problem is that school reform is a large-scale issue, and REAR is virtually ignorant about large-scale processes in education. Researchers study small-scale phenomena for very short periods of time. Their knowledge comes from controlled laboratory research, pilot

studies, case studies in a few schools, or a few examples of unusually effective schools. Newer research techniques, such as meta-analysis, have been developed that "pretend" that outcomes were generated on a large scale. REAR then convinces itself and the profession that the knowledge gained from these small-scale investigations is applicable.

Reality 7. It's the scale, stupid! Large-scale change reflects properties that are often diametrically opposed to those in effect in small-scale research. While small-scale success is inspirational, the methods are not necessarily workable on a large scale. The fact that something works in a few classrooms, in a few schools, with a few teachers, at a few grade levels, for a few weeks, and so on says nothing about whether or how it can be disseminated or will actually work on a large scale. Conversely, the fact that a standardized test is inaccurate in the case of a few individuals doesn't mean that it isn't giving an accurate overall picture of large-scale results.

There are almost no cases in which researchers have studied an innovative practice on a large scale, as it was happening, over an extended period of time. Indeed, there aren't even appropriate research methods in education for evaluating large-scale effectiveness.¹¹ Thus reforms supposedly based on research are, at best, little more than hunches that are usually based on inapplicable studies.¹²

As a profession we do not even ask the right questions about the large-scale efficacy of ideas. For example, instead of asking whether full inclusion works, the large-scale point of view would ask whether there exists a sufficiently articulated technology to allow 80% of the sites that exert a reasonable effort actually to make full inclusion work. If full inclusion or any other reform cannot be made to work consistently, we should not attempt to make it national policy. Rather, we should treat such reforms as ideas about which we can learn from those who voluntarily choose to pursue them. Alternative assessment is currently an innovative idea, not a technology or an innovation.

Myth 8. Directive programs cannot be effective on a large scale, and attempts to implement such programs rob teachers of their individuality. This myth denies the whole history of the performing arts.

Reality 8. It is possible to develop a

new generation of far more powerful programs that can be effective on a large scale. The performing arts have survived and flourished because they have been able to systematize the delivery of highly creative performances by striking an appropriate balance between directive components — e.g., scripts and choreography — and individual interpretation. The same thing can be done in education. The success of HOTS and other creative programs, such as Reading Recovery and Junior Great Books, suggests that it is possible to develop programs that combine the best of educational research and pedagogy into specified systems that consistently generate high levels of learning and also stimulate highly creative forms of interaction between students and teachers on a large scale.

Consequences of the Myths

For the last 100 years REAR has been using the myths described above to develop and promote highly amorphous and ill-advised reforms. In the absence of substantive methodologies, these reforms are of necessity implemented through highly simplistic strategies that are doomed to fail. For example, full inclusion is implemented by eliminating pullouts and treating everyone the same — instead of by developing programs to meet the needs of all students. The nature of the reforms produced by these myths leads to a number of consequences.

Repeated failure of reform initiatives. The failure of reforms produces much trauma for practitioners and students. Indeed, reforms based on the above myths fail not only once, but often a number of times. For example, consider the child-centered school. In one incarnation, the advocacy of child-centeredness is associated with the use of middle schools as an alternative to the content orientation of junior high schools. However, at the turn of the century, the goal of child-centeredness was a major rationale for the creation of a new type of school to be called the “junior high school” — as an alternative to the content-oriented high school. Unfortunately, we do not know how to create child-centered schools that increase learning and social development on a reasonably consistent basis. But fear not! REAR will create another type of school that starts a grade lower and will be child-

centered for sure.

Massive waste of resources on staff development and dissemination. In the absence of a valid technology and a body of experience for carrying out a proposed reform on a large scale, inservice training and dissemination strategies are largely ritualistic shams that waste time. Staff development gets everyone excited, but little happens. Still, despite its woeful track record, we keep on pushing staff development with religious fervor as *the* key to improving education.

No professional validation standards for considering and implementing reform.

We have no tradition of insisting on anything approaching reasonable validation of proposed reforms before we rush to implementation. This lack became glaringly evident to me some time ago as I was watching TV. Jonas Salk was announcing that recent work in his lab had convinced him that AIDS could be cured by a particular type of vaccine. Instead of calling for the production of vaccine, he went on to say that he needed to conduct large-scale tests over the next five to 10 years to figure out how viable the use and production of the vaccine was. Then the news switched to some education policy maker who was advocating the national adoption of the latest reform proposals for school improvement. The two news clips captured the difference in professional responsibility between medicine and education, and we looked silly in comparison.

Seymour Sarason expresses the same idea as follows:

To a significant degree, the major educational problems stem from the fact that educators not only accepted responsibility for schooling but, more fateful, also adopted a stance that essentially said: we know how to solve and manage the problems of schooling in America. Educators did not say: there is much we do not know. . . .

[T]he medical community has made a virtue of its ignorance insofar as its stance with the public is concerned. That community did not say that it would be able to cure cancers next year or 20 years from now. On the contrary, it emphasized the complexity and scope of the problems, the inadequacies of past and present conceptions and practices....¹³

A current example of this lemming-like approach to reform is site-based management. The site-based bandwagon is at

least five years old, but there still is no agreement on what it is and no evidence that it is either workable or effective.¹⁴ So why is everyone doing site-based management and so many other current reforms? Indeed, why is the current administrative training literature so focused on leadership for change with almost no standards for the conditions under which administrators should resist change? Anyone standing pat is automatically seen as reactionary.

Repeated cycling of inadequate progressive and traditional reforms. The absence of adequate validation norms for proposed reforms has meant that the primary reason to adopt reforms is the failure of what exists rather than the demonstrated success of what is being proposed. For example, in a recent article, Diane Ravitch noted that “the educational results bear the reformers out.” I read on eagerly, thinking I was going to find out about improved educational results from some new reform. Unfortunately, her rationale was a litany of the failures of the status quo: “Fewer than half the city’s ninth-graders graduate within four years. Of those who do, nearly 40% enter the City University of New York, and only a quarter of those pass all three of its tests of minimal reading, writing, and math skills.”¹⁵ Similarly, curriculum theorists point to the repeated failure of nationally generated curricula to argue that curricula should be developed locally — even though there is no evidence offered that locally developed curricula work better.¹⁶

As legitimate problems with the status quo become apparent, the profession periodically tries progressive ideas. When the progressive reforms fail, the traditionalists come back with their own set of inadequate proposals. Neither side ever seems to get it right. This yo-yo effect has been going on for more than a century and is worse than a broken clock; at least the clock is correct twice a day.

Nowhere is the yo-yo syndrome truer than in Title I. Over the years results have consistently found little effect beyond the third grade. In the next-to-last reauthorization, the traditionalists prevailed, and strict basic skills accountability guidelines were incorporated into the legislation. This had little effect. In the latest reauthorization, progressives moved to the fore, and the legislation emphasized schoolwide reform. This probably won’t work either. In

**Meta-analysis
is a technique to
make small-scale
research seem to
be large scale.**

truth, neither side had large-scale experience in improving learning for educationally disadvantaged students. The sides merely waged political war on the basis of their philosophies and myths. A whole series of fundamental changes that were critical for real improvement were not even considered.¹⁷

Misleading conclusions and misleading uses of research. When researchers have little experience with the phenomena that they seek to apply their findings to, the result is usually misleading conclusions. That is not to say that researchers deliberately mislead the field. Rather, they mislead one another, which then causes them inadvertently to mislead the field.

Indeed, one of the reasons why HOTS is so effective is that I was ignorant of the research on the development of thinking. In particular, I was ignorant about the fact that nearly everyone agreed that thinking was best developed in content. Thus I didn't think twice when my experience suggested that harnessing the tremendous intellectual potential of educationally disadvantaged students in grades 4 through 7 required first developing their sense of understanding through immersion in general thinking activities divorced from classroom content.¹⁸ Unfortunately, researchers have no instincts by which to judge whether the conclusions in the literature make sense.¹⁹

It is impossible to develop a true understanding of the nature of student/teacher interaction from reading research. Indeed, as I was writing this article, the then-latest issue of the prestigious *American Educational Research Journal (AERJ)* arrived. It featured a special section on fostering higher-order thinking skills. I im-

mediately put aside my writing to see whether it had any significant new ideas that could help improve HOTS.

The first of the articles found that teachers were much more likely to ask higher-order thinking questions of higher-performing students than of lower-performing ones.²⁰ Surprise! In their tortured attempt to explain why this was so, the authors blamed tests, institutional norms, teachers, and so on — but they ignored the obvious possibility that lower-performing students might not in fact know how to respond to such questions.

In the second article, researchers followed one first-grade teacher for four years to learn how she used the research-based knowledge about student learning that they had provided to guide her teaching of addition and subtraction. The researchers concluded that the teacher was teaching differently from the way they would have expected her to. As a result, the authors concluded that “teachers would be better translators of knowledge about children’s thinking than we would.”²¹ Unfortunately, the researchers never did document that the teacher was in fact applying research at all.

The third article studied 12 sixth-graders in two science classrooms. Each student was observed for the equivalent of four lessons to determine why some students understand science concepts while others do not. The researchers solemnly concluded that the differences in engagement and understanding they observed in these students were due to “complex interactions among cognitive qualities of academic tasks, students’ knowledge and achievement, and students’ motivational and affective orientations in science classrooms.” So what else is new?²² The *AERJ* research was disappointing, silly, misleading, and useless, and it was predicated on the validity of the above myths.

Even more misleading have been the research results from meta-analyses. Meta-analysis is a relatively new technique to make small-scale research seem to be large scale. In meta-analysis, results from a series of individual small-scale studies are aggregated into a single finding. This single finding appears statistically to have been generated from a single large-scale study.

Much of the push for full inclusion and schoolwide models of reform has come from a series of meta-analyses that find small effects from special programs, such

as gifted programs.²³ Unfortunately, meta-analyses combine results from a series of studies in which virtually nothing is known about the nature or quality of the interventions that generated the results. Indeed, I have even seen a case in which a researcher generated meta-analysis results for the effects of a program that he was unaware had not been implemented and did not exist. Since the researchers are almost always ignorant about the quality of the programs in the included studies, meta-analysis crunches data generated from primarily weak programs and thereby severely underestimates the impact of effective programs.²⁴

Substitution of philosophy of process for philosophy of outcomes, of good intentions for science, and of global good efforts for precise interventions. Faced with the repeated failure of reforms, we have essentially shifted from a focus on outcomes to a focus on processes. We choose up sides based on the process of a proposed approach. People say such things as “I am philosophically opposed to pull-outs” or “I am philosophically opposed to prescriptive programs” or “I am philosophically opposed to phonics.” When you cannot produce the outcomes you need, all that’s left is to look good philosophically while you fail.

We have also abandoned science and cognition in the design of reforms. Saying “All children can learn” says nothing about *what* they can learn, *how fast* different youngsters can learn, *when* different types of children are best able to learn certain things and why, and so on. Or consider the reform breakthrough on how best to use a computer — “Use it as a tool!” That advice says nothing about what to build, how to build it, and what other tools you might need.

In the absence of specific, systematic interventions that work, reformers become obsessed with getting everyone on board, infusing a reform throughout the curriculum, and carrying out the process all the time. Untracking comes to mean that students should *never* be separated according to ability (except for sports, of course). Opposition to the mindless use of standardized tests or rote learning or pull-out programs comes to mean that these technologies should *never* be used.

Yet, despite all these problems, the myths persist. The ultimate reality is that the only way to improve education sig-

nificantly is by the use of more powerful forms of curricula in the hands of very good teachers who are trained to teach better. All three of these conditions must exist. Any other type of reform is a sham — no matter how compelling its philosophical rationale. Any proposed reform should be tested against whether it is likely to enhance these three conditions consistently and directly.

Why Do Practitioners Tolerate The Ignorance of REAR?

The short answer is that the activities of the REsearch/Academic/Reform community provide hope. Practitioners passionately, even desperately, want to help young people. The hardest thing for a teacher or principal to accept is failure with a student who has potential and needs help. Good practitioners are always seeking answers and searching for something that will help them be more successful. Unfortunately, the answers provided by REAR are more often than not illusory. If you remove yourself from the fray for a moment to sit back and think about it, building movements around such concepts as school restructuring, using the computer as a tool, or creating child-centered schools is silly. But in our desire to help students, we do not sit back.

REAR also provides excuses for failure. It spews out ever more esoteric, jargon-ridden philosophical rationales to explain why apparent failures aren't really failures, why it's impossible to be more systematic, and why a new reform is needed. REAR's intellectual machinations help shield the profession from outside criticism, while also providing rallying cries and rationales for new funding.

Some of the current defense mechanisms being deployed by REAR include blaming society, blaming tests, or claiming that schools are as good as they ever were. Gerald Bracey's 1994 report in this journal promotes the excuse that students who do poorly do so because of demographic factors, such as poverty, that are beyond the control of schools.²⁵ This is like arguing that we shouldn't expect to be able to fly because gravity is beyond human control. Given that we have spent billions trying to get schools to perform better for economically disadvantaged students, inadequate progress should not be excused.

Beginning Anew: Reforming the Reformers

Instead of leading, REAR is inhibiting substantive improvement. The members of the REAR community currently act like spoiled siblings who cannot get along. It is time for them to grow up and stop viewing education as a playing field for their ideologies. We need to stop the silly posturing about knowing what works when there is no proof that any of these reforms work on even a small scale. This rhetoric has outlived its usefulness and has become a self-delusional detriment to educational progress. The dissemination of research knowledge and inservice training as the *primary* vehicles for reform has failed and is unworkable.

The sad thing is that many progressive ideas have great potential, if only the technology appropriate for them had been developed. For example, my staff and I recently employed constructivism to develop a unique solution to a classic problem: teaching and learning word problems in math. No, we did not create a national network of courses shot through with hype about the use of constructivism. Rather, we devised a new technology for helping students create their own sense of how language and math go together. We developed a new type of software to enable students to communicate with a lonesome space creature stuck inside their computer — a creature that understands English and speaks math. The result of using this software, along with appropriate teacher interaction, was that students came to view word problems on tests as simple.²⁶

Creating better techniques and technologies requires increasing investment in development, slowing the rush to large-scale implementation, and rethinking the role and structure of colleges of education. We no longer need colleges composed largely of individuals and courses that spread the latest incarnations of unworkable myths. Rather, we need organizations that can integrate research and philosophy with the development and large-scale testing of new technologies. Such organizations would have fewer courses and far more joint-development ventures involving university faculty members, students, and practitioners. Teachers and students could work together to design interventions and collect data on their effectiveness. This would force faculty members to confront the limita-

tions of their ideas and subject them to review by those who must implement them.

One problem might arise in that colleges of education do not have the highly skilled inventors, craftspeople, and tinkerers who can put the pieces together. There are few individuals with expertise in such areas as biology, artificial intelligence, or the integration of social services, which are likely to be critical for generating new technologies. Moreover, very few people in colleges of education have experience in designing technologies. Converting ideas into successful innovations will require new types of knowledge generators, disseminators, and reformers. We need to increase the diversity of skills within education faculties, much as we have increased racial and sexual diversity. What exists in education today is the equivalent of having only theoretical physicists and philosophers on the faculty of an engineering school. I wonder what a bridge built by theoretical physicists and philosophers would look like?

A vibrant research and reform community is critical to the future of education. Unfortunately, for much of this century, these communities have been in a rut. They have been able to generate many innovative ideas but few innovations — that is, effective large-scale reforms.

Education can no longer afford a research and academic community that is detached from the real processes that take place in schools and from the large-scale consequences of the ideas that it proposes. Education can no longer afford a well-intentioned but inept progressive movement and a too-limited traditional movement, each waiting for its 10- to 20-year turn in the limelight. It can no longer afford the piling on of ill-conceived movements to compensate for the inadequacy of current ideas. The result has been inefficiency and waste.

We need to start by being honest and saying that we do *not* know what works for the educationally disadvantaged student, that we do *not* know how to get most students thinking on a higher level, that we do *not* know how to increase their motivation to learn, and that we do *not* know how to systematically blend the best of progressive and traditional ideas. We also need to recognize that caring and empowerment are not enough and that movements are no substitutes for better tech-

niques. Significantly improving the learning of educationally disadvantaged students on a large scale requires fundamental breakthroughs in the development of powerful and highly systematic technologies, and people must be willing to invest a decade of work in the pursuit of that goal. Finally, we need to learn that mass advocacy should follow, not precede, the careful development and large-scale testing of techniques.

It is not easy in any field to develop interventions that avoid prevalent myths, are creative, and meet the criteria set forth by Drucker. However, experience has shown that it is possible to design interventions for education that are practical for large-scale use, that are highly creative, that incorporate many progressive ideals, and that consistently produce higher levels of traditional outcomes. Producing such interventions requires some new skills and some different conceptions of what it means to be a researcher, a reformer, or a practitioner. A few academicians and reformers have taken up the challenge. Far more are needed.

1. For information about the HOTS program, see Stanley Pogrow, "Making Reform Work for the Educationally Disadvantaged," *Educational Leadership*, February 1995, pp. 20-24; idem, *HOTS (Higher Order Thinking Skills): A Validated Thinking Skills Approach to Using Computers with Students Who Are At-Risk* (New York: Scholastic, 1990); idem, "A Socratic Approach to Using Computers with At-Risk Students," *Educational Leadership*, February 1990, pp. 61-67; and idem, "Challenging At-Risk Students: Findings from the HOTS Program," *Phi Delta Kappan*, January 1990, pp. 389-97.

2. For discussions of which types of innovations survive, see David Tyack, Michael Kirst, and Elisabeth Hansot, *Educational Reform: Retrospect and Prospect* (Palo Alto, Calif.: Institute for Research on Educational Finance and Governance, Stanford University, Project Report #79-A5, September 1979); and David Tyack and William Tobin, "The Grammar of Schooling: Why Has It Been So Hard to Change?," *American Educational Research Journal*, Fall 1994, pp. 453-79. For a discussion of what happens to innovations that survive, see Larry Cuban, "What Happens to Reforms That Last? The Case of the Junior High School," *American Educational Research Journal*, Summer 1992, pp. 227-51; and idem, "The Lure of Curriculum Reform and Its Pitiful History," *Phi Delta Kappan*, October 1993, pp. 182-85.

3. For the failure of state-level reform, see Robert A. Frahm, "The Failure of Connecticut's Reform Plan: Lessons for the Nation," *Phi Delta Kappan*, October 1994, pp. 156-59; and Judith McQuaide and Ann-Maureen Pliska, "The Challenge to Pennsylvania's Education Reform," *Educational Leadership*, December 1993/January 1994, pp. 16-21. For the failure of school restructuring and site-based management, see Jeffrey Mirel, "School Reform Un-

plugged: The Bensenville New American School Project, 1991-93," *American Educational Research Journal*, Fall 1994, pp. 481-518; and Dianne Taylor and Ira E. Bogotch, "School-Level Effects of Teachers' Participation in Decision Making," *Educational Evaluation and Policy Analysis*, Fall 1994, pp. 302-19.

4. Peter F. Drucker, *Innovation and Entrepreneurship* (New York: Harper & Row, 1985).

5. Stanley Pogrow, "Where's the Beef? Looking for Exemplary Materials," *Educational Leadership*, May 1993, pp. 39-45.

6. Carl F. Kaestle, "The Awful Reputation of Educational Research," *Educational Researcher*, January/February 1993, pp. 23-31.

7. Claude Goldenberg and Ronald Gallimore, "Local Knowledge, Research Knowledge, and Educational Change: A Case Study of Early Spanish Reading Improvement," *Educational Researcher*, November 1991, pp. 2-14.

8. Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

9. My research has shown that you can stimulate the development of thinking skills in educationally disadvantaged students and so spark a wide variety of improvements in learning in just 35 minutes a day of exemplary activities over a two-year period. For a discussion of this 35-minute principle, see Pogrow, "Challenging At-Risk Students"; and idem, "Converting At-Risk Students into Reflective Learners," in Arthur L. Costa, James A. Bellanca, and Robin Fogarty, eds., *If Minds Matter: A Foreword to the Future* (Palatine, Ill.: Skylight Publishing, 1992).

10. In some content fields I could not find expert academicians who were willing to look at curricula and make judgments about the relative quality of the materials; in others, the judges panicked when the moment came to apply the criteria that they had developed. Academicians seemed more interested in making relativistic arguments as to why ratings could not be done.

11. It makes no practical sense to use comparison groups to determine the large-scale effectiveness of an intervention. A more appropriate statistical procedure would be to determine the consistency of effects across 50 to 100 treatment sites and simply to assume that, if consistent effects are occurring, it is a result of the intervention. However, such a study would probably not be published in the top research journals.

12. Here's a personal example of how researchers misapply small-scale findings. Researchers tell me that research has found that engaging students in discussions on the use of computers reduces their enjoyment of the technology and so reduces its potential for learning. Yet Socratic dialogue about computer experiences is the key element in the success of the HOTS program. While the finding cited by researchers is true over the short term, my large-scale experience has consistently found that, after several months of discussing computer use, students exhibit far higher levels of cognitive development and enjoyment than the use of technology alone could generate.

13. Seymour B. Sarason, *The Predictable Failure of Educational Reform* (San Francisco: Jossey-Bass, 1990), pp. 37-38.

14. The fall 1994 issue of *Educational Evaluation and Policy Analysis* featured a special section on site-based management/shared decision making. Two of the four articles were devoted to aspects of

conceptualizing what site-based management is. The third article studied the process of implementing shared decision making in six schools and found heightened conflict with little effect on school reform. The fourth article went beyond process and studied the effects in one district of shared decision making on student achievement and teacher practice. (It was good to find a researcher actually interested in student learning.) The fundamental conclusion was that student achievement had not improved and that teachers had not changed their instructional practices.

15. Diane Ravitch, "First, Save the Schools," *New York Times*, 27 June 1994, p. A-17.

16. For examples of those recommending local curricula, see Larry Cuban, "The Lure of Curricular Reform and Its Pitiful History"; and Decker Walker, *Fundamentals of Curriculum* (New York: Harcourt Brace Jovanovich, 1990), pp. 307-36.

17. For alternative policy recommendations for Chapter 1/Title I, see Stanley Pogrow, "The Forgotten Question in the Chapter 1 Debate: Why Are the Students Having So Much Trouble Learning?," *Education Week*, 26 May 1993, pp. 36, 26; and idem, "What to Do About Chapter 1: An Alternative View from the Street," *Phi Delta Kappan*, April 1992, pp. 624-30.

18. See the theory of cognitive underpinnings in Pogrow, "Challenging At-Risk Students"; and idem, "Converting At-Risk Students."

19. When it became clear that HOTS was producing significant gains, I went back to the research literature to see where the conclusion that thinking should first be developed in content had originated. Almost all the research had been conducted with highly educated adults as subjects; it had nothing to do with educationally disadvantaged students in middle school.

20. Steven Raudenbush, Brian Rowan, and Yuk-Fai Cheong, "Higher Order Instructional Goals in Secondary Schools: Class, Teacher, and School Influences," *American Educational Research Journal*, Fall 1993, pp. 523-54.

21. Elizabeth Fennema et al., "Using Children's Mathematical Knowledge in Instruction," *American Educational Research Journal*, Fall 1993, pp. 555-84.

22. Okhee Lee and Charles W. Anderson, "Task Engagement and Conceptual Change in Middle School Science Classrooms," *American Educational Research Journal*, Fall 1993, pp. 585-610.

23. Edward Baker, Margaret Wang, and Herbert Walberg, "The Effects of Inclusion on Learning," *Educational Leadership*, December 1994/January 1995, pp. 33-35.

24. For a criticism of the tendency of the meta-analyses conducted by Robert Slavin to overlook the effects of high-quality programs and thereby to grossly underestimate the effects of good programs, see Susan Demirsky Allan, "Ability Grouping Research Reviews: What Do They Say About Grouping and the Gifted?," *Educational Leadership*, March 1991, pp. 60-65; and Stanley Pogrow, "Good Statistics About Bad Programs Tell Little," *Educational Leadership*, letter to the editor, October 1991, p. 93.

25. Gerald W. Bracey, "The Fourth Bracey Report on the Condition of Public Education," *Phi Delta Kappan*, October 1994, pp. 114-27.

26. The software is called Word Problem Processors. For more information on this software, contact the author by fax at 520/621-9373.