

# On Research and School Organization: A Conversation with Bob Slavin

At the Johns Hopkins Center for Research on Elementary and Middle Schools, Bob Slavin directs the elementary school program and monitors research findings on a broad scale.

Here he explains Success for All, an early intervention that organizes resources to ensure that students arrive at 3rd grade with adequate skills, and calls for practitioners to make stringent demands upon researchers for experimental trials of instructional techniques.

**O**ver the last year or so you've looked at a lot of research on various ways to organize schools. Why?

I am broadly interested in what makes a difference in achievement, particularly in the elementary grades. When I run up against an issue, I do a thorough review of the literature so that I am confident about what the research actually says. Before I write it up for others, I first have to convince myself.

**How do you do it? Do you personally search out the original research reports, analyze them, give more weight to some studies than others, and so on? Is this "meta-analysis"?**

I have my own method, which I call "best evidence synthesis." It owes a good deal to meta-analysis but pays more attention to the quality of the studies—to substantive issues rather than only statistical issues—in deciding what to include and what not to include.

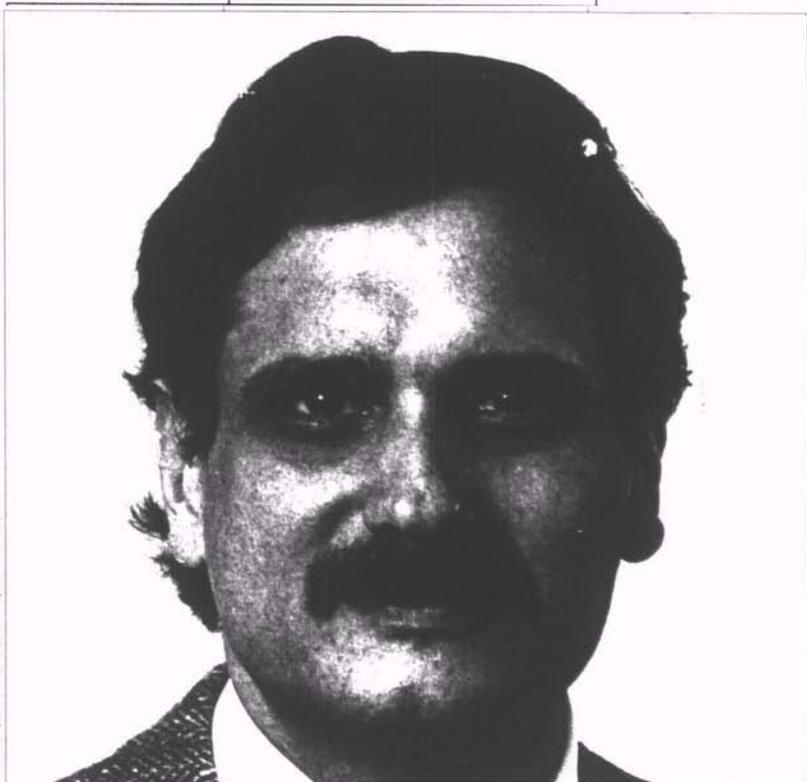
**One topic you looked at this way was ability grouping (see Slavin 1988). What are some others?**

Mastery learning, class size, cooperative learning. Those are the ones on which I have done complete, formal best evidence syntheses.

**When you put all your findings together, what sort of model**

**of school organization do they suggest?**

To make schools markedly more effective, they really have to be quite different from the way they are now, particularly in serving students who have the greatest difficulty. It takes a three-part strategy to meet the needs of kids in difficulty. First and most important is much more emphasis on prevention: on picking up on problems very early and making sure that kids don't get into the special education or the remedial cycle. Second, we must change instructional methods to use systematic programs that have clear evidence of effectiveness. Only after we implement effective preventive programs and classroom change programs should we be looking at the



third issue, which is remediation. As one example of this kind of approach, we are currently doing research on a program that uses one-to-one tutoring, a completely restructured reading curriculum, family support services, and other components to try to make sure that kids in an inner-city school don't arrive at 3rd grade without adequate skills.

**So you have in fact devised a comprehensive model?**

Yes, it is called "Success for All."

**And it reflects what you have learned from looking at the research on grouping, class size, tutoring, and so on?**

Exactly. It doesn't just load it all in, of course. The various components have to work together in a coherent and meaningful way—but you want to be on firm ground as you make each choice.

**What are some characteristics of the model that reflect the research on school organization?**

One is the use of tutoring, which really came out of looking at class size research. A lot of people ask whether there isn't a critical point at which class size begins to make a difference, and the answer is yes, but—that critical class size is one student. You don't really see a lot of difference even when classes are very, very small, but one-to-

one tutoring is vastly more effective than even one-to-five teaching.

That finally led us to try to get away from the Chapter I approach of pulling out small groups, a mild treatment that may go on throughout the student's entire elementary career. Instead we decided to invest our resources in one-to-one tutoring that would get kids up to grade level rapidly and then try to maintain them within the regular classroom.

Another aspect of the program is a form of the Joplin Plan: grouping students across grade lines according to reading levels. First, 2nd and 3rd graders who are all performing at the 2-1 level may be in a class together. We do this because, as I said in my review

(see Slavin 1988), the kind of grouping that has the greatest research support is cross-grade grouping.

We also have evidence from other sources about the uselessness of the workbook activities that students usually do in reading classes. Then we have the time-on-task literature and some of the direct instruction literature, which suggest that it is beneficial to increase the amount of time that students work actively with the teacher.

We also provide extended day and preschool programs because of the strong evidence for effectiveness of preschool and extended kindergarten for disadvantaged kids. In these programs we use the Peabody Language Development Kits, which also have strong research support.

**How long will it be before you know whether this model is effective?**

We now have data for one entire school year—we also assessed the students every eight weeks to check on their progress, to regroup, and to decide who needed tutoring—and we have seen remarkable progress. Our promise is that a kid who starts with us in preschool will get to 3rd grade with adequate skills; that no child will fail unless he or she is severely retarded. It may be three years before we know whether we can fulfill that promise, but we know that mean scores have increased this year.

**We decided to invest our resources in one-to-one tutoring that would get kids up to grade level rapidly.**

**Is the program quite a bit more expensive?**

Oh yes. We are proposing a shift of Chapter I monies to earlier grades so there will not be so much remediation to do in the upper grades. We already know that we can reduce special education referrals, so we may be able to demonstrate cost-effectiveness even before the students get to the middle school level.

**Let me turn to another topic. One of the more controversial conclusions you reached in your reviews of the research literature is that mastery learning is not nearly as effective as claimed (see Slavin 1987). How did you decide that?**

It is actually very simple. I did a number of things in my review, but if this were all I did, it would have been enough: I required that the studies had to have taken place over at least four weeks. The studies that produced the big effects—the ones that Bloom talks about and that are cited in a lot of other mastery learning syntheses—were conducted in three days, one week, two weeks, three weeks. Requiring that the treatment had to be in place for at least four weeks brings down the mastery learning studies to a very small number, and I think that even four weeks is really too short.

Another attribute I looked for was use of a control group. I think a control group is essential, because a lot of people evaluated mastery learning programs during the period 1979-1985, which was also a time when people were discovering all kinds of other ways to increase their test scores: curriculum alignment, retaining students in grade, things like that. When scores were reported for control groups within the same district, then everybody was operating under the same ground rules—how the tests were given and so on. But in studies with control groups the difference between the control groups and the mastery learning groups was practically nonexistent.

**But for many educators it seems pointless to argue with mastery learning. Surely we should try to**

**Our promise is that a kid who starts with us in preschool will get to 3rd grade with adequate skills.**

**be clear about the outcomes we want, to assess student learning periodically, and not to go to the next level of difficulty until students have the prerequisites. It seems self-evident.**

There are lots of principles in education that are correct but that don't result in educational gain. The concept of mastery learning is almost axiomatically true, but the issue is what it means in actual practice. I am talking here only about group-based mastery learning; if you have a 10-day lesson, you teach for eight days and then give a test. The kids who need correctives get corrective instruction, while the other kids get enrichment activities. What may be going on is that two days of corrective instruction are not enough for kids who have serious problems, but two days are too much wasted time for kids who don't have a problem. In other words, Bloom's Learning for Mastery model may actually be too limited a response to student differences in learning rates.

**So it's not that it's too sophisticated; it may not be sophisticated enough. If you are going to do something different from traditional instruction, why not do something better?**

Precisely. I think it does not go far enough to ensure students' success. We use mastery learning principles in our Success for All model, but we use

them more directly. The purpose of the tutor is not to teach something different, but to ensure that each child has learned what the classroom teacher has taught. If a teacher teaches long "a" in the morning, the main task of the tutor is to be sure the students get long "a," so that tomorrow they are ready for the next lesson. That's clearly a mastery learning concept, but we do it daily and with a one-to-one tutor to make sure that the kids actually get it.

**But as you say, it's in accord with the principles of mastery learning. I doubt that Benjamin Bloom and other advocates would quarrel with your design.**

Yes, but another problem with some mastery programs is that they chop knowledge into little pieces, and that's unnecessary. What is critical about mastery learning or any other program is how it's done. Another good example is individualized instruction. Nobody would quarrel with giving kids instruction suited to their needs; but in the '60s and early '70s, when educators made individualized materials and had the kids work on them all by themselves, the effect on achievement was disastrous. The concept was correct, but the way in which it was operationalized was foolish, because while you were accomplishing one goal—adapting instruction to individual needs—you were interfering with other goals, such as providing students with explanations by quali-

**Bloom's Learning for Mastery model may actually be too limited a response to student differences in learning rates.**

fied teachers and motivating students to learn.

**Some of the things that you say don't work have been advertised as based on research. How would you describe the state of educational research these days?**

It's not at the stage I wish it were, but in the last 15 years educational research has come into its own. We have a lot of promising developments: for example, the process-product studies of effective teaching. I am encouraged by the emphasis on what's really going on in classrooms.

## Response to Slavin: Mastery Learning Works

James Block

This year marks the 20th birthday of mastery learning. In schooling, concepts that last even a few years are powerful and rare. As such, it is not surprising that mastery should draw periodic potshots. Slavin's is but the most recent.

I am disturbed by Slavin's interpretation of the cumulative mastery learning research record. Mastery advocates (e.g., Guskey and Gates 1986, Guskey and Pigott 1988), as well as more neutral parties (Stallings and Stipek 1986, Walberg 1985), have repeatedly reported exactly what Slavin proposes—"experimental-control comparisons over realistic periods of time with suitable measures of effectiveness." The most recent review, in fact, involves some 83 quasi-longitudinal studies ranging from years to over a decade (see Block et al. 1988). These studies have concluded, almost without exception, that mastery works. The basic strategies move the student who achieves at the 50th percentile to somewhere between the 65th and the 86th percentiles; this movement occurs in ordinary elementary and secondary school subjects; and the change registers on well-known standardized tests (Block et al. 1988). Enhanced mastery learning approaches do even better (Spady and Jones 1985).

I am also saddened by the dampening effects that this interpretation may have on communication between mastery learning and "Success for All" practitioners. There is a striking resemblance between Slavin's agenda for improving schooling and mastery learning's (see Block et al. 1988, pp. 49-57). Slavin's interest in the development of component approaches to instruction parallels mastery's promotion of the development of functional ones. His interest in the prevention of student learning difficulties jibes with mastery's, too. His interest in advancing "cooperative" learning overlaps with mastery's interest in developing "self-determined" learners. Lastly, his interest in tutoring clearly correlates with mastery's interest in the development of talent.

So I hope Slavin will not dismiss these comments out-of-hand. They are meant to caution him, his followers, and fair-minded observers not to throw out the proverbial mastery learning baby with the bathwater. That baby is now a young adult. They are also meant to suggest to Slavin and his followers that collaboration rather than confrontation is now in order. There are good reasons for the vitality and longevity of the mastery concept. While Slavin waits three years for data on his ideas, we are willing to share our 20 years of experience now.

Block, J., H. Eftim, and R. Burns. *Building Effective Mastery Learning Schools*. New York: Longman, 1988.

Guskey, T., and S. Gates. "Synthesis of Research on the Effects of Mastery Learning in Elementary and Secondary School Classrooms." *Educational Leadership* 43, 8 (1986): 72-80.

Guskey, T., and T. Pigott. "Research on Group-Based Mastery Learning Programs: A Meta-Analysis." *Journal of Educational Research* 81, 4 (1988): 198-216.

Spady, W., and B. F. Jones. "Enhanced Mastery Learning and Quality of Instruction." In *Improving Student Achievement through Mastery Learning Programs*, edited by D. U. Levine. San Francisco, Calif.: Jossey-Bass, 1985.

Stallings, J., and D. Stipek. "Research on Early Childhood and Elementary School Teaching Programs." In *Handbook of Research on Teaching*, 3rd ed., edited by M. C. Wittrock. New York: Macmillan, 1986.

Walberg, H. "Examining the Theory, Practice, and Outcomes of Mastery Learning." In *Improving Student Achievement through Mastery Learning Programs*, edited by D. U. Levine. San Francisco, Calif.: Jossey-Bass, 1985.

James Block is Professor of Education, University of California, Santa Barbara, CA 93106.

## We use mastery learning principles in our Success for All model, but we use them more directly.

### Let's clarify the term *process-product*. What does it refer to?

Studies in which a large group of teachers are carefully observed and the teachers whose students made the greatest gains are compared with those whose student gains were much smaller. A lot of the teacher behaviors people are talking about now in terms of classroom management, components of the lesson, and so on are direct outgrowths of that approach.

### What are the advantages of that kind of research?

One is that you don't go in with preconceived notions that you try to impose on teachers. You are focusing on practices you know are realistic, because they already exist in classrooms. That is also a limitation, of course; you won't discover anything very surprising, because the only practices you will see will be those in fairly widespread use.

There are two major strands of research that contribute directly to practice. One is *correlational* research—of which the process-product research is a good example—where researchers identify the best of current practice. The other is *experimental*, where researchers try to increase the range of variation by introducing new methods

that a teacher might never stumble on otherwise. My own belief is that until you've done experimental research, you shouldn't be talking to teachers at all. I feel that a lot of the correlational and descriptive research is valuable in *suggesting* what kinds of things we should look at, but that it should never be used as a guide to practice.

### Why do you say that?

I think that in order to tell teachers, "Do this and your students will learn better," you must have compared the results of that kind of advice using

control and experimental groups. For example, if you studied basketball players and found that the taller ones did better, you might then say, "Well, we've learned that taller basketball players do better, so let's start stretching our shorter players." The correlational finding is correct; it is a useful thing to know. But it doesn't mean that a stretched player is the same as one who is naturally tall.

Similarly, with teachers we often find things from correlational research that are truly characteristic of the best teachers—but that doesn't mean that

## Response to Slavin: Who Defines Best?

Thomas R. Guskey

Robert Slavin is a respected educational researcher who has made a number of valuable contributions. Yet his recent work involving what he labels "best-evidence" syntheses and his application of this technique have evoked strong criticism from many within the research community. "Best-evidence" synthesis procedures were proposed by Slavin (1986) as a way to focus more attention on the quality of the studies being synthesized. But what has become apparent is that "best," in this particular sense, is a relative term (Joyce 1987) and that the best-evidence syntheses Slavin conducts include *only* those studies that he, and he alone, considers best (Gamoran 1987, Hiebert 1987). The results of these best-evidence syntheses, therefore, are often potentially biased, highly subjective, and likely to be misleading (Guskey 1987).

A case in point is Slavin's best-evidence synthesis of the research on mastery learning (Slavin 1987). In this review Slavin threw out all studies of less than four weeks' duration, even though previous reviews of mastery learning research that have specifically investigated the influence of study duration have consistently shown it has no effect on study results (Block and Burns 1976, Guskey and Gates 1985). He also threw out all studies which compared the results teachers attained before-to-after their implementation of mastery learning and studies involving multiple teachers, each teaching different subjects. For a significant portion of the review, he threw out all studies in which teachers used class time to help students with their corrective work and included only those studies in which correctives were done outside of class as independent work. In addition, he rejected all studies that did not use results of standardized, norm-referenced achievement tests as a principal outcome, even though mastery learning has always been based on the premise that student learning should be evaluated in terms of criterion-referenced standards. In the end, he had only seven studies, looked at their results, and concluded that mastery learning doesn't work. But if you narrowly restrict the way in which an innovation is implemented, alter a basic component of that innovation, and then evaluate its effectiveness in terms of an outcome poorly aligned with its intended goals, why would you expect positive results? What is most unfortunate, however, is that while many people now know this was Slavin's conclusion, very few understand the rather idiosyncratic process by which that conclusion was reached.

There are, at the same time, several unusual ironies in all this. Some years ago Benjamin Bloom (1984a, b) and his graduate students (Anania 1981, Burke 1983) conducted a series of studies emphasizing the especially powerful effects of tutoring. The model that Slavin now describes emphasizes the use of tutoring. Also, Slavin

those are the best things to teach to other teachers. The process-product people admit, at least in private, that a lot of the individual teacher behaviors that they have found to be characteristic of effective teachers are really just indicators of hard-working, intelligent, insightful, very active teachers. Now, how do you get other teachers to exhibit these behaviors? That's a very different question.

**It seems reasonable to try to teach those same behaviors to other teachers.**

But in actual fact many of the programs derived directly from the process-product research have failed. One clear example is the Madeline Hunter programs. Everything Madeline Hunter talks about is very well established, either in laboratory research or correlational research or both; but when assembled into an experimental program, will it work in actual practice? We now have enough evidence to say that it doesn't.

**What makes you say that? A lot of people think it works very well.**

now criticizes mastery learning because some programs "chop knowledge into little pieces, and that's unnecessary." Yet his program would insist students master such discrete skills as "long 'a'." A final irony: in 1981, in *All Our Children Learning*, Bloom outlined his ideas on a comprehensive model of schooling, which considered the characteristics of teachers, students, and their interaction; the curriculum and instructional materials; peer groups; and the home environment. Slavin's comprehensive model, which includes "one-to-one tutoring, a completely restructured curriculum, family support services, and other components" is labeled "Success for All." Researchers hope, as descriptions of this model are developed, they will include appropriate citations.

- Anania, J. "The Effects of Quality of Instruction on the Cognitive and Affective Learning of Students." Doctoral diss., University of Chicago, 1981.
- Block, J. H., and R. B. Burns. "Mastery Learning." In *Review of Research in Education*, Vol. 4, edited by L. S. Shulman. Itasca, Ill.: Peacock, 1976.
- Bloom, B. S. *All Our Children Learning*. New York: McGraw-Hill, 1981.
- Bloom, B. S. "The Search for Methods of Group Instruction As Effective As One-to-One Tutoring." *Educational Leadership* 41, 8 (1984a): 4-18.
- Bloom, B. S. "The 2-Sigma Problem: The Search for Methods of Group Instruction As Effective As One-to-One Tutoring." *Educational Researcher* 13, 6 (1984b): 4-16.
- Burke, A. J. "Students' Potential for Learning Contrasted Under Tutorial and Group Approaches to Instruction." Doctoral diss., University of Chicago, 1983.
- Gamoran, A. "Organization, Instruction, and the Effects of Ability Grouping: Comment on Slavin's 'Best-Evidence Synthesis.'" *Review of Educational Research* 57 (1987): 341-345.
- Guskey, T. R. "Rethinking Mastery Learning Reconsidered." *Review of Educational Research* 57 (1987): 225-229.
- Guskey, T. R., and S. L. Gates. "A Synthesis of Research on Group-Based Mastery Learning Programs." Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1985.
- Hiebert, E. H. "The Context of Instruction and Student Learning: An Examination of Slavin's Assumptions." *Review of Educational Research* 57 (1987): 337-340.
- Joyce, B. "A Rigid Yet Delicate Touch: A Response to Slavin's Proposal for 'Best-Evidence' Reviews." *Educational Researcher* 16, 4 (1987): 12-14.
- Slavin, R. E. "Best-Evidence Synthesis: An Alternative to Meta-Analysis and Traditional Reviews." *Educational Researcher* 15, 9 (1986): 5-11.
- Slavin, R. E. "Mastery Learning Reconsidered." *Review of Educational Research* 57 (1987): 175-213.

**Thomas R. Guskey** is Professor, Department of Educational Policy Studies and Evaluation, College of Education, University of Kentucky, Lexington, KY 40506.

I don't think they would if they looked at the research. The study that Jane Stallings did in Napa, California (Stallings and Krasavage 1986), a study in West Orange, New Jersey (Donovan et al. 1987); a study involving the entire state of South Carolina (Mandeville 1988)—all three found essentially no difference between control classes and those using Hunter's methods.

**Aren't there other explanations for the results in each of those instances?**

Even with the problems that people cite, if there had been big effects, researchers would have found something, but in South Carolina, for example, where 15,000 teachers were trained in the Hunter method, researchers found nothing at all.

**You're saying that in rushing to devise programs based on correlational research, educators may have jumped the gun a bit.**

I'd be delighted if people were rushing out to design programs based on that research and then were carefully evaluating them. But I think it's

**Research is never going to inform practice as it should until practitioners demand top quality experimental evidence before they adopt programs and practices on a large scale.**

### Response to Slavin: Toward a Greater Variety

Benjamin S. Bloom

When used correctly, mastery learning does improve student learning. However, no single method of improving learning will work well for all teachers, all school subjects, and all students.

We must have a greater variety of methods and programs to ensure good learning for most of the students in our schools. I am delighted that Robert Slavin recognizes this.

In a recent paper, I described needed improvements in mastery learning, initial prerequisites in a new course or subject, the support of the home environment, developing automaticity in reading, and teaching the higher mental processes (Bloom 1988).

Perhaps at some time in the distant future, even Robert Slavin will declare a truce in his opposition to mastery learning.

Bloom, B. S. "Helping All Children Learn Well in Elementary School—and Beyond." *Principal* 67, 4 (March 1988): 12-17.

Benjamin S. Bloom is Charles H. Swift Distinguished Service Professor, Department of Education, University of Chicago, 5835 Kimbark Ave., Chicago, IL 60637.

One response to what you're saying is that much of what we're doing right now doesn't have any validity so we have to keep trying what appears to be better even if we don't have the data.

You have to use your best judgment, yes. But we must demand better data for making important education decisions.

The other problem, of course, is that researchers argue among themselves about what, in fact, is true. One of the best-accepted tenets in recent years has been that mastery learning has powerful effects—and now you tell us it isn't so.

It would certainly be helpful if we had widely agreed upon standards for evaluation and more third-party agencies that were conducting evaluations. It would help us get away from some

not appropriate to design such programs on the basis that "correlational research supports this, so it has to work in the classroom, and that's my proof."

### What does that mean for the future of educational research?

In my view, research is never going to inform practice as it should until practitioners demand top quality experimental evidence before they adopt programs and practices on a large scale. They will also have to conduct, or allow to be conducted, good experimental studies of new methods within their own districts when they try them on a small scale. Until we have practitioners who say, "I'm not going to use that until I've

seen good experimental evidence for it," we're going to continue on the educational pendulum with the "miracle of the month" and will not make much serious progress.

**Until you've done experimental research, you shouldn't be talking to teachers at all.**

### Response to Slavin: What's the Best Evidence?

Herbert J. Walberg

Robert Slavin does me great honor by criticizing me along with Benjamin Bloom, who ranks with Gary Becker, James Coleman, Piaget, Carl Rogers, Skinner, and Margaret Wang among this century's great educational theorists. Credit for mastery learning research and synthesis, however, should hardly be accorded to me, since my part has been to calculate and compile estimates of the learning effects of many successful and popular educational interventions, including mastery learning (Walberg 1984).

The best recent estimate of mastery effects is large and close to Bloom's estimate and the one I cite (Kulik and Kulik 1986). To conclude that more time yields more learning or that short-term mastery studies show bigger effects than longer ones requires comprehensive evidence—dozens of studies of mastery—not Slavin's hand-picked "best evidence." The same may be said for other interventions for which we now have substantial evidence: some 8,000 studies and 112 reviews of them in my latest collection.

Slavin makes an expansive and unfulfilled promise; he claims "remarkable progress" with "Success for All" despite his apparently incomplete, uncontrolled, and unpublished single study. Yet he calls for "third-party" evaluation. Interested readers will find such evaluation of Slavin's work on mastery learning, cooperative learning, ability grouping, and class size in the highly critical commentaries on his theories, methods, and conclusions by about a dozen researchers that have followed his recent articles in the *Educational Researcher* and *Review of Educational Research*.

Kulik, J., and C. L. Kulik. "Mastery Testing and Student Learning." *Journal of Educational Technology Systems* 15 (1986): 325-345.

Walberg, H. J. "Improving the Productivity of America's Schools." *Educational Leadership* 41 (May 1984): 19-27.

Herbert J. Walberg is an Educational Psychologist and Research Professor of Education, University of Illinois at Chicago, 522 N. Euclid Ave., Oak Park, IL 60302.

## Response to Slavin: Improving Teacher Decisions

Madeline Hunter

I heartily commend Robert Slavin for his endeavors to take a research-based analytic look at what can make a difference in education. I also appreciate the time he has spent becoming familiar with our teacher decision-making model since he wrote "The Hunterization of American Schools," where he characterized it as "rigid and mechanistic." Now he correctly states that the model is based on research that is "well established either in laboratory research or correlational research or both," before he claims the evidence shows it "won't work in actual practice." Let's examine that evidence.

- The Napa Project, given its flaws, presented hard data that test scores escalated when teachers were applying what they had learned. What the project really validated was that newly learned professional skills do not maintain themselves without encouragement and refinement from coaching. In addition, the teaching "skills" were used only for reading and math; therefore the teachers did not perceive them as generic elements in every teaching decision regardless of content or teaching mode. I criticized the project for the omission of the transfer theory that is central to the model.

- Manatt reports positive results with the SIM Project. An associate and I were active in establishing the integrity of content and the certification of its transfer from information to knowledge to judgment to wisdom, for both teachers and administrators. Without judgment and wisdom there is a danger of teachers' becoming robots.

- Although I have been involved in some workshops in West Orange, New Jersey, I cannot attest to the integrity of use of the model. I have serious questions about the general rigidity and "assembly line procedure" of some of the New Jersey Academy's work.

- In terms of the South Carolina study, the researchers state, "The unfortunate truth is that higher-level coaching skills are crucial to teacher improvement." Those skills were missing in many of the coaches. This condition also prevailed in the third year of the Napa Project. The amazing result in both projects was that teachers' attitudes were very positive.

I have cautioned educators in South Carolina against both the mass production of courses and the use of trainers with little or no certification of competence or content integrity. There I also cautioned the state superintendent against rigidity, lack of decision-making skills, and the adoption of a checklist. This fall I will be working with trainers in South Carolina and Arkansas to build correction into the "assembly line" aspects that unfortunately have developed.

Slavin ignores the original validation of the model (Project Linkage) conducted by an independent investigator, Rodney Skagar, which substantiated impressive gains by students in an inner-city Los Angeles school where the integrity of the model was certified. Also, how does he explain the 20-year escalation of acceptance all over the world? Surely educators are not that gullible—they must be seeing results.

There is no aerodynamic research available to support, and there is much to refute, a bumblebee's flying. But the bee flies! If practitioners really said, "I'm not going to use that until I've seen good experimental evidence for it," our classrooms would be immobilized. What research does Slavin believe is supporting current practice?

The Hunter model is an effort to change teaching decisions, many of which are based on tradition, folklore, and fantasy, to theory-based judgments and wisdom. Only by so changing can we move classroom teaching into behaviors that more closely approximate the success of tutoring, something we'll never be able to finance and which, with expert teaching, will seldom be needed.

Madeline Hunter is Professor of Education, University of California—Los Angeles, Department of Education, Moore Hall, 405 Hilgard Ave., Los Angeles, CA 90024-1521.

craziness had no bearing on what they do on a day-to-day basis. Because we don't have well recognized standards for what constitutes an adequate evaluation, Walberg took the approach of taking the studies that had "mastery learning" in the title, putting them all together, and reporting a number for them.

### Your advice to practitioners then is to look more carefully at the data.

Yes. Before you accept the next program that comes down the pike, demand experimental-control comparisons in real schools over realistic periods of time with suitable measures of effectiveness.

### You're saying the evidence is out there?

Yes. And there'll be a lot more—if people start demanding it. □

### References

- Donovan, J. F., D. A. Sousa, and H. J. Walberg. "The Impact of Staff Development on Implementation and Student Achievement." *Journal of Educational Research* 80 (1987): 348-351.
- Mandeville, G. K. "An Evaluation of PET Using Extant Achievement Data." Paper presented at the annual convention of the American Educational Research Association, New Orleans, April 1988.
- Slavin, R. E. "Mastery Learning Reconsidered." *Review of Educational Research* 57 (1987): 175-213.
- Slavin, R. E. "Synthesis of Research on Grouping in Elementary and Secondary Schools." *Educational Leadership* 46 (September 1988): 67-77.
- Stallings, J., and E. M. Krasavage. "Program Implementation and Student Achievement in a Four-Year Madeline Hunter Follow-Through Project." *Elementary School Journal* 87 (1986): 117-138.
- Walberg, H. J. "Improving the Productivity of America's Schools." *Educational Leadership* 41 (May 1984): 19-27.

**Robert E. Slavin** is Director, Elementary School Program, Center for Research on Elementary and Middle Schools, The Johns Hopkins University, 3505 N. Charles St., Baltimore, MD 21218. **Ron Brandt** is ASCD's Executive Editor.

of the arguments among researchers about what constitutes adequate evidence. If, for example, you told practitioners clearly that much of the evidence on mastery learning that Herb Walberg (1984) summarized in his

synthesis was taken from studies of one week's duration, in which the corrective instruction, which the control group did not receive, was given in addition to regular class time, they would know immediately that such

Copyright © 1988 by the Association for Supervision and Curriculum Development. All rights reserved.