The Journal of Educational Thought, Vol. 12, No. 3 (1978)

REVIEW ESSAYS

The review deals with two central themes of *Teachers Make a Difference:* the claim expressed in that title, and the assertion that researchers can help make an even greater difference in educational outcomes. The book presents syntheses of research findings, notably from correlational studies showing that elementary teachers do have a differential influence on cognitive outcomes. Yet in stressing statistical relationships, the authors skirt the issue of substantive and practical significance. They do not pinpoint ways in which research might identify useful behaviors which might be taught to teachers. Rather, they tend simply to assume that once policy makers decide on goals, researchers can find means to implement them in rational fashion.

Robert Donmoyer
Greg Farman
Michael Imber
Cheryl Perry
Beverly Showers
Daniel L. Duke

Do Researchers Make a Difference?

Good, Thomas L., Biddle, Bruce J., and Brophy, Jere E. *Teachers Make a Difference.* New York: Holt, Rinehart and Winston, 1975. 271 pp., $5.95

Teachers Make a Difference is a work whose title only partially characterizes its contents but conceals its authors' multiple intentions. *Researchers Make a Difference* would have provided an equally informative title, since the book is as much an argument for teacher effectiveness research as it is for the effectiveness of teachers.

Although these themes — that teachers make a difference and that researchers can help teachers make an even greater difference — comprise the central elements of the book, a number of other topics also are included. The second half of the book is devoted largely to discussions of the characteristics and comparative advantages of norm-referenced and criterion-referenced test, the strengths and limitations of behavioral objectives, the growth and meaning of the accountability movement, and the value to teachers of doing research on their own effectiveness. The authors also attempt to define and describe "open education" and "individualization" and to review the literature relevant to both. Finally, evidence is cited to substantiate two generalizations which those concerned with teaching should find useful: (1) students with different characteristics probably need different instructional treatments if they are to make maximum progress; (2) even teacher behaviors which are effective may become ineffective if overused.

Unfortunately, the book's central themes are not strong enough to unify the mass of information described above. The presence of so much information, in fact, tends to obscure the authors' basic arguments. Still, most of the information is presented with clarity and objectivity, and its inclusion undoubtedly enhances the utility of the book for both preservice and inservice teacher education.

The arguments that teachers and researchers make a difference are clearly the authors' central concerns and also the source of most of the book's problems. Therefore, this review will concentrate on these two themes. The discussion will be organized around four questions. First, have the authors provided sufficient evidence to support their conclusion that teachers make a difference? Second, how accurately does the authors' conception of making a difference reflect the goals and purposes usually associated with schooling? Third, how sound are the authors' arguments for continuing and expanding educational research? Finally, is the process-product model which undergirds the authors' commitment to educational research an adequate representation of reality?

Do Teachers Make a Difference?

Good, Biddle, and Brophy do not wish to make the relatively trivial argument that teachers (or schools) make a difference when compared with no teachers (or no schools). Rather, their argument is that some teachers (and some schools) make *more* of a difference than other teachers (and schools).[1] Although they cite several studies to reinforce the point, the authors concede early on that "(m)ost would agree that schools make a difference in the broadest sense. Students enrolled in a school do better than students not attending any school (p. 2)."

The authors employ a two-fold strategy to argue for the existence of differential teacher effects. First they attack the conclusions of researchers who contend that differential outcomes of schools are merely the result of different student entering characteristics rather than differences in school environments. The authors argue, for example, that in the data used by Coleman and reanalyzed by Jencks, "... school quality is so tied up with community and pupil characteristics that an independent measure of it cannot be obtained (p. 22)." They go on to contend that even if the data from Coleman and Jencks could demonstrate that schools exert no differential effects, the data would provide no information about differential *teacher* effectiveness since the unit of analysis employed by Coleman and Jencks was the school and not the classroom. Assuming that effective teachers are distributed randomly, Good, Biddle, and Brophy maintain that the effects of these teachers would not have been reflected in the data. In the same vein, the conclusions of J. Stephens receive criticism along with the position of Arthur Jensen and others who have made assumptions about the effectiveness of schooling on the basis of I.Q. data. Although much of the discussion of the I.Q. controversy seems irrelevant to the argument for teacher effectiveness, the authors do successfully execute the first part of their strategy. Particularly in their discussion of the data used by Coleman and Jencks, the authors demonstrate that the question of differential teacher effectiveness is still open for discussion.

Having shown that the skeptics' contentions concerning the effectiveness of schools and teachers are flawed, the authors proceed to the second part of their argument. As evidence for the existence of differential effects, they present a rather extensive review of the literature on the effectiveness not only of teachers, but also of schools and curricula. Large scale studies of teacher effectiveness, especially the work of Robert Soar and Brophy and Evertson, receive particular attention.[2] Concerning the research design employed by Brophy and Evertson in their attempt to identify particularly effective teaching behaviors, the authors write,

> The sample selection process by itself provided evidence that teachers differ significantly from one another in the amount of learning their students accomplish. About half of the teachers showed consistent patterns from one year to the next, and among these, some teachers were consistently quite successful in producing student learning gains while others were consistently unsuccessful. . .The data clearly showed that teachers have differential effects on student learning which are both statistically and practically significant (p. 75-76).

The conclusion that certain teachers (and certain schools and curricula) are more effective than others appears convincing. The authors present an impressive array of nontechnical summaries, focusing on diverse and previously unlinked research findings to establish the existence of differential effects. Their presentation is a welcome antidote to the earlier work of Coleman, et al., and may do much to cure the professional malaise brought on by it. Still, the evidence presented by Good, Biddle, and Brophy is subject to certain limitations.

One problem centers on the fact that most of the evidence comes from research on elementary schools. The authors acknowledge that there are not many studies relating teaching behaviors with learning gains at the secondary level. They cite only Flanders' review of several studies utilizing his classroom observation system and two additional studies suggesting correlations between school or classroom climate variables and student achievement. Since the handful of studies presented is almost exclusively correlational and since the correlations are usually low, the evidence presented to establish the existence of differential teacher effects at the secondary level is hardly compelling.

A second problem is that the evidence cited by Good, Biddle, and Brophy to support their proposition that teachers produce differential effects centers on a limited range of effects. Although a handful of studies are cited to demonstrate differential teacher effects on affective outcomes and such variables are creativity and student morale, the bulk of the evidence focuses on those types of cognitive learning commonly measured by standardized achievement tests. Even when affective variables are considered, there is a tendency to view them as independent variables which might affect students' scores on standardized achievement tests rather than as ends in themselves.[3]

When the authors speak of both non-cognitive outcomes and cognitive outcomes at the secondary level, they assume that citing a few positive findings is sufficient to establish the existence of differential teacher effects. Yet correlational studies of social phenomena always will be incapable of establishing causality, and experimental studies which attempt to control the complexity of real classrooms inevitably will call forth questions of task validity.[4] The best that can be hoped for in educational research is that a preponderance of evidence will make inferences about causality convincing. The authors' assertions about non-cognitive outcomes and cognitive outcomes at the secondary level are clearly hypotheses, not established facts; neither the volume nor the character of evidence provided in support of these hypotheses justifies the authors' assertion that teachers make a difference in these particular areas.

In summary Good, Biddle, and Brophy appear to demonstrate that teachers produce some differential effects, particularly in cognitive learning at the elementary level. There remains, however, a problem more basic than the authors' blanket contention that teachers make a difference. The discussion now turns to a consideration of this problem.

What Kind of Difference Do Teachers Make?

Statisticians rightly have cautioned against confusing statistical significance with substantive significance, yet this is precisely what Good, Biddle, and Brophy do.[5] Although they contend that the differential effects produced by some teachers are "both statistically and practically significant (p. 76)," the evidence they present supports only statistical significance. They do not show how statistical significance relates to practical significance. In fact, the substantive sense in which teachers supposedly make a difference is never discussed in the book.

It is misleading to assume that a teacher who succeeds in getting his students to produce test results which are significantly higher than average in a statistical sense has accomplished anything of practical importance. Whether a third grader score at a 3.5 or a 4.5 level on a standardized reading or math test may or may not be practically important. Practical significance can only be discussed if it can be demonstrated or if it is assumed that higher scores are indicative of something else that is valued. The authors make no attempt to refute recent research which fails to find a correlation between standardized achievement test scores and success in later life.[6]

The authors largely ignore those substantive differences which have historically been associated with schooling. Consider, for example, the assumption that schooling can eliminate poverty, an assumption which dates back at least to Horace Mann and has been widespread ever since.[7] Although the authors indicate an awareness of this historical concern in their chapter on "Goals for Education," they make no attempt to relate it to their discussion of differential effects. The efforts of Jencks and others to demonstrate that the traditional assumption about schooling and inequality is fallacious are overlooked. The authors ignore other acknowledged substantive goals such as socialization, social reform, and character development which have traditionally been associated with schooling.

It could be argued that questions of substantive significance are too complex ever to be answered by educational researchers. Good, Biddle, and Brophy, however, take quite a different position. In keeping with their unfailing faith in the power of research, they state that, "the more complex the goals taken on by the school, the more we need the help of researchers if we are to accomplish these goals (p. 88)." It is this unfounded faith in the power of research which will now be considered.

Do Researchers Make a Difference?

The second major focus of *Teachers Make a Difference* centers on the tremendous potential which the book's authors see for educational research. It is the authors' belief that, ultimately, through research, education can and should become a vastly improved applied science. On the basis of this belief, they make several pleas for an expansion of research and for increased funding of the type of projects they advocate. While it seems clear that there is a great need for continuing education research based on carefully conceived methodologies, nevertheless, it must be noted that there are problems both with the authors' expectations for teacher effectiveness research in particular and with their beliefs concerning the potential of educational research in general.

The authors argue that it is possible for researchers to identify effective teaching behaviors through classroom observations of exemplary teachers. They tend, however, to oversell existing research findings and to underestimate the limitations of the process-product conception of research.

One problem is that virtually all of the findings offered as evidence that effective teaching behaviors can be identified are correlational rather than experimental. As noted in the previous section, causation cannot be established by such studies. Furthermore, the authors admit that "these teaching behaviors do not correlate very strongly with student outcome measures (p. 58)." They attempt to account for the fact that "the strength of the relationship revealed in these studies is not impressive (p. 58)" by arguing that student entering characteristics and other uncontrolled variables were not taken into account. However, these same contextual variables lead Lee Cronbach to a very different conclusion. On the basis of his extensive work with Aptitude X Treatment interactions, he presents a much more humble appraisal of the potential benefits of research in education. Because of higher order interactions, Cronbach contends, it is unlikely that research will ever be able to derive the type of generalizations which Good, Biddle, and Brophy promise. Cronbach writes,

> Systematic inquiry can realistically hope to make two contributions. One reasonable aspiration is to assess local events accurately, to improve short run control. . .The other reasonable aspiration is to develop explanatory concepts, concepts that will help people to use their heads. . . Though from persistent work in many contexts he (the researcher) may reach an actuarial generalization of some power, this will rarely be a basis for direct control of any single operation.[8,9]

Even if casual relationships between specific teaching behaviors and student outcomes could be generalized, the utility of such research findings would be questionable. Many of the teacher traits which the authors cite as positively correlated with student learning outcomes, for example, may lack the sort of operational clarity which would make them useful for the preservice and inservice education of teachers. Although the authors criticize others for using umbrella terms such as "indirect teaching" which have a variety of different, perhaps unrelated, meanings, many of the teacher qualities they suggest as possibly related to effectiveness are no more precise. Therefore, it is not at all clear whether teachers would be able to develop these qualities, or whether teacher educators would be able to recognize and reinforce them, much less whether they would produce the predicted results.[10]

In addition to overselling the benefits of teacher effectiveness research, Good, Biddle, and Brophy underestimate the difficulties which process-product research must surmount in order to address some of the foremost concerns of education. The authors fail to note the near impossibility of fitting many of the ideas and values which are central to education into their approach. Heubner has pointed out that aesthetic and ethical concerns do not fit readily into a means-ends framework.[11] Even when concerns can be conceptualized in terms of educational outcomes, there are often practical problems with measurement which cannot be overcome as easily as the authors would lead their readers to believe. Perhaps they are correct in asserting that progress has been made in measuring non-cognitive and higher order cognitive outcomes and that increased effort will net even more refined measuring instruments. However, measuring affective variables and such hard to define concepts as creativity and higher order cognitive outcomes undoubtedly will continue to require the use of constructs which only grossly approximate reality. The authors themselves demonstrate the distortion that can occur when the fit between an operational definition and a substantive meaning is loose. On the basis of several studies, they report that open education does not foster creativity, yet the operational definition of creativity used in the studies, scores on the Torrence Test, is hardly an adequate representation of the multitude of meanings associated with that term.[12]

Thus, the authors' argument that researchers make a difference is hardly compelling. It ignores the liabilities and limitations of process-product research, while depending for positive evidence on studies that are less impressive than the authors would contend. The studies cited by Good, Biddle, and Brophy fail to demonstrate that researchers will be able to provide the kinds of generalizations they prophesy. Even if causal generalizations were forthcoming, the utility of such information for the training and retraining of teachers has yet to be demonstrated.

There is, however, a more basic problem with the authors' conception of the role that researchers can play in educational policy making. This problem stems from the authors' belief in the rational model of educational decision making. The adequacy of this model can be questioned on both logical and empirical grounds.[13]

Does the Rational Model Provide a Satisfactory Representation of Reality?

Good, Biddle, and Brophy's commitment to the rational model of educational decision making can be clearly seen in their chapter entitled, "Goals For Education." In this chapter the authors echo Max Weber's position that means and ends are separable. They call for "national efforts to clarify educational goals and to provide information about how such goals can be realized (p. 99)." The assumption is that questions of educational ends are value issues which must be resolved through the political process. Once such issues are resolved, however, the authors believe that the researcher can determine the most efficient means of reaching those ends.

The assumption that means and ends can be separated for anything other than analytical purposes is open to question. Joseph Schwab, in discussing the process of curriulum decision making, indicates that both ends and means must be constantly considered and decision makers "must treat them as mutually determining one another."[14] Charles Lindblom adopts a similar position with respect to policy making in general, arguing that means and ends can never be separated, since values can only be discovered through the process of making decisions about specific means.[15]

According to the view of Good, Biddle, and Brophy, a community could determine *a priori* that its primary educational goal was the cognitive achievement of its children, while leaving to the experts the technical question of how this goal might be realized. Researchers might then suggest that the most effective way to reach this goal would be to segregate students on the basis of standardized test scores. However, the analysis of Lindblom and Schwab shows that it is wrong to think that this finding could be implemented automatically by responsible decision makers. Rather, it would have to be considered in light of a number of potentially undesirable consequences, such as racial segregation.

Studies of organizational decision making provide empirical evidence that the rational model's assumption about the separation of means and ends is not a satisfactory representation of reality.[16] Recent unsuccessful attempts to implement PPBS provide additional empirical support for the proposition that the rational model is unrealistic.[17]

The educational researcher's role in the decision making process must, therefore, be quite different from the one envisioned by the authors. The researcher can facilitate a well focused debate by providing technical information and helping to conceptualize issues, but he can never hope to provide pat answers to problems which are politically and socially as well as educationally complex.

Conclusion

This review has focused on two central themes of *Teachers Make a Difference.* It was noted that Good, Biddle, and Brophy have demonstrated that elementary school teachers produce differential effects in certain areas measured by standardized achievement tests and that these effects are statistically significant. The unimpressive evidence that secondary school teachers "make a difference" and the authors' failure to distinguish between statistical and substantive significance were noted.

The argument that researchers make a difference was not found to be convincing. Although recent research in the area of teacher effectiveness has produced interesting and potentially useful results, the assertion that such research can identify particularly effective teaching behaviors which can be taught to teachers and would-be teachers has yet to be validated. Effectiveness may well be a transient phenomenon. What is effective for one teacher at one time may not be effective for the same individual at another time. Finally, the authors were criticized for ignoring important limitations of process-product research in particular and the rational model in general.

Despite these criticisms, notice must be taken of Good, Biddle, and Brophy's contribution to the field of teacher education. They synthesize considerable research data and develop a strong case for more rigorous, multi-dimensional investigations of teacher effects. The best potential effect of their book, however, is that it may give a much needed morale boost to a beleaguered profession.

Notes

[1] The authors also note in passing that some teachers are effective at some times but not at others.

[2] Unfortunately, the teacher effectiveness data from the California Commission study (MacDonald) were not yet available when *Teachers Make a Difference* was published. The findings from this large scale research enterprise serve to bolster some of the authors' contentions.

[3] Thomas L. Good, Bruce J. Biddle, and Jere E. Brophy, *Teachers Make a Difference* (New York: Holt, Rinehart, and Winston, 1975), pp. 72-74.

[4] Lee Shulman, "Reconstruction of Educational Research," *The Review of Educational Research,* 40 (1970), 371-395.

[5] Leslie Kish, *Survey Sampling* (New York: John Wiley and Sons, 1967), pp. 587-589.

[6] Christopher Jencks and Marsha D. Brown, "Effects of High Schools on Their Students," *Harvard Educational Review,* 45 (August 1975).

[7] Horace Mann, *The Republic and the School,* ed. Lawrence A. Cremin (New York: Teachers College Press, 1957).

[8] Lee Cronback, "Beyond the Two Disciplines of Scientific Psychology," *American Psychologist,* 30 (February 1975).

[9] Michael Scriven makes a similar point. He notes that even the natural scientist cannot provide the type of predictive power required of the psychological and educational researcher. The physicist, for example, can no more predict how a specific leaf will fall from a specific tree on a specific autumn day than a psychological researcher can predict the effects of a specific educational treatment on a specific student. The difference between the physicist and the psychological researcher, Scriven points out, is that no one is particularly interested in predicting the behavior of a leaf and therefore the variables about which the physicist has no knowledge and cannot control do not interfere with his work. The same thing cannot be said of the psychological and educational researcher. Michael Scriven, "Views of Human Nature," *Behaviorism and Phenomenology,* ed. T. W. Wann (Chicago: The University of Chicago Press, 1964).

[10] N.L. Gage and his associates at Stanford University are currently exploring some of these issues.

[11] Dwayne Heubner, "Curricular Language and Classroom Meanings," *Curriculum Theorizing,* ed. William Pinar (Berkeley: McCutchan Publishing Co., 1975), pp. 217-236.

[12] Robert Stake has indicated he believes "that sometimes it would be preferable to evaluate the quality of the opportunity to learn, the 'intrinsic' merit of the experience rather than the more elusive 'payoff'..." Robert Stake, *Evaluating the Arts in Education* (Columbus, Ohio: Charles E. Merrill Publishing Company, 1975).

[13] For a discussion of the rational decision making model see Graham T. Allison, Chapter 1, *Essence of Decision: Explaining the Cuban Missile Crisis*(Boston: Little, Brown, 1971).

[14] Joseph Schwab, "The Practical: A Language for Curriculum," *School Review,*(November 1969), pp. 1-23.

[15] Charles Lindblom, *The Intelligence of Democracy*(New York: Macmillan, 1965).

[16] Graham T. Allison, Chapter 1, *Essence of Decision: Explaining the Cuban Missile Crisis* (Boston: Little, Brown, 1971).

[17] Michael Kirst, "The Rise and Fall of PPBS in California," *Phi Delta Kappan* 56 (April 1975), pp. 535-538.