



On Evaluation: An Interview with Daniel L. Stufflebeam

Ronald Brandt

Daniel L. Stufflebeam is Director of the Evaluation Center at Western Michigan University, Kalamazoo. According to Stufflebeam, "Evaluation has to be more than an academic exercise. Unless there are important decisions to be made, we ought not to be wasting money on it."

Brandt: There seems to be increasing dissatisfaction among some teachers and principals about the amount of testing and data gathering that goes on in schools. Other people claim there is too little information on which to base decisions, which suggests that we should be gathering even more. How can those two positions be resolved?

Stufflebeam: We often gather too much information or the wrong information. For example, we don't need all the standardized testing data that we have. When I was holding hearings on the Michigan Accountability System, the representatives from Detroit came in with the standardized tests then being used in the Detroit Public Schools. The stack was about two feet tall. They said they didn't need another test; they

needed systematic help in problem solving behavior.

Their point was that they were getting too much of the wrong information. The State Assessment Tests in Michigan were administered in the spring and the results would finally get back to teachers in November or December. By that time the information was cold and irrelevant.

Brandt: How do you feel about the idea that we should stop using standardized tests altogether because they are so bad?

Stufflebeam: It isn't the tests that are bad; it's our use of them. We've gotten into a kind of mindless perseverance in the administration of standardized tests and the reporting out of percentile ranks for students in schools and school dis-

tricts. Very little constructive use is being made of the results. The main use is to embarrass school districts that rank low and to pat others on the back for a sterling performance. As we know, that performance is usually correlated with the socioeconomic levels of students.

If we are serious about evaluating for improvement purposes and as a guide to decision making, we need to identify our purposes and the kinds of decisions to be made, and then formulate and utilize evaluation procedures that make sense. If we do it that way, we will find that the standardized testing program is often not very useful.

Brandt: What are the alternatives?

Stufflebeam: Many people are dissatisfied with the empha-

sis on cognitive development. There is tension in society because the high school diploma is required for entry into adult life, but many students who rate low on indicators of cognitive development are able to perform vocationally and economically. The society has to decide whether to let people into adulthood with low academic credentials or to withhold the high school diploma from them.

One aspect of the problem is related to the variables we measure. As we have said, we have identified one area of development—the cognitive. But it seems to me that schools do much more than that. There are other variables that are terribly important, and we ought to give students credit for development in those areas. We ought to give ourselves as educators credit for what we're doing in those areas.

The ones I have thought about fall in seven categories: In addition to cognitive development, there are social development, physical development, affective development, aesthetic development, vocational development, and moral development. Those seven areas might be reflected in requirements for graduation and in reports on student progress that are sent home to parents.

Brandt: How realistic is it to think we could devise performance measures for those things?

Stufflebeam: I don't know for sure. I know that consider-

able research and development have been going on in all of them, and I am trying to gather the best of what has been done. It may be a better way to grant school credit than the way we usually do it.

Brandt: You have had a lot to do with shaping educational evaluation in the last decade. In the late 1960s, you often spoke about the poor quality of evaluation. How do you feel about the state of the art now?

Stufflebeam: Evaluation was a new idea for us back in the 1960s. In practice we hadn't done much of it, and what had been done followed one of two main models. Some—mostly individual research projects—used the experimental design approach. We hadn't had enough experience to know that under real world conditions it would not work the way it was supposed to.

The other way was the accreditation approach. We thought there wasn't time to do a true experiment, so we would bring in a group and let them use professional judgment.

When the federal government began forcing more evaluations, we tried to use those two approaches. They didn't work. It was a waste of money, and it wasn't giving usable answers.

In the last few years, we've had lots of turmoil, lots of ideas. Evaluation still isn't as good as it ought to be, but we're finding better approaches and beginning to have a desirable impact.

Brandt: Your idea of evaluation includes much more than looking at the outcome or product of a program. You also include the processes being employed, the planning of those processes, and even the original goals. How does an evaluator go about evaluating goals? Isn't that just a matter of values?

Stufflebeam: First of all, it is very important that it be done. Misdirected activities often waste large sums of money and do disservice to people. We have to make sure that our goals are justified and responsive to the needs of the people to be served.

The best way to do it is on a prospective, proactive basis. To me that means conducting an ongoing needs assessment of the students we are serving. It means that before we arbitrarily set off in a certain direction, we have thought about who we are serving, what we want them to be able to do, what they want to be able to do, and what they can and cannot do. We have some data to use in setting our objectives.

Brandt: One of the problems faced by school districts is that the broad goals they have adopted seem so unrelated to what really goes on in classrooms. What can be done to bring about a better match between district goals and instructional programs?

Stufflebeam: Goal setting is too often an abstract activity that has no reality base. We

convene to write goals and objectives for students we have not seen and about whom we have no information. Then, when we encounter the students and learn of their needs, we forget about the goals. In many cases, it is just as well to forget about the written goals, because they are simply rhetoric.

We are more likely not to forget the goals if they are set on the basis of a needs assessment. Then we can use the goals to guide and assess instruction.

They won't be forgotten if they are used as criteria for deciding such things as what textbooks to buy. Goals should be based on identified needs and then used as criteria for evaluating program plans.

A Look at Needs Assessment

Brandt: You have referred several times to "needs assessment." There are a lot of different ideas about what that is. What is yours?

Stufflebeam: I could describe some steps that I think ought to occur in a needs assessment. I wouldn't claim that this is original. I've only tried to synthesize the good elements in needs assessments I've looked at.

The first step is to establish the purpose for the needs assessment. Often there are relevant policies that are at issue. An example might be to provide equal educational opportunities to minority groups. In addition to the purpose, we need to iden-



tify the audience to be served and the subjects of the evaluation.

Planning an assessment of needs forces one to think about what a need is, so I went to the dictionary to find out. A need is "a variable that is requisite or necessary, desirable or useful" in regard to satisfying that policy we mentioned. So, step two is to identify variables of interest—variables that are pertinent to the policies, the audience, and the subjects.

Now we must distinguish between met and unmet needs. That's a key distinction, because if we focus just on the unmet needs, we may fail in the future to meet needs that are currently being met. A met need is the possession of something that is necessary, desirable, or useful for implementing the policy of interest, and an unmet need is the lack of whatever

that is. That may help us think about the kinds of variables we are looking for.

Some of the variables we think of will be dependent variables, and others will be independent. The policy may spell out some dependent variables—some outcomes we hope to achieve. An example is equal educational opportunity for all students. We can also think of potential side effects that must be considered, and we want those on the right side of our equation too.

As for independent variables, we have learned some things about teaching and learning and about organizing and administering schools. So we should look to research and theory to help us identify independent variables that would affect implementation of the policy. Of course, we also do the needs assessment in a politi-



cal situation, so we should not look only in books. We have to ask the key powers in the situation what they think some of the important variables are. We compile a list of variables based on what the research tells us and what the people in the situation think are important.

All right, we've got a set of variables. Next we operationalize them in terms of data collection approaches, which may include use of instruments. It will also be useful later on if we can get people to establish or estimate critical levels on those instruments once we have identified them. All of this, remember, is still a part of identification of variables.

Once we have the instruments and the critical levels, we move to step three, which is getting judgments about importance of each of the variables. The reason for that is that later on, we may find that we have

25 different needs to be investigated, but that we can't afford to investigate them all. If they have been ranked for importance beforehand, we will have a better handle on how to assign priorities to needs and objectives. I would get various reference groups—teachers, parents, students, administrators, board members, community members—to give their ideas of the relative importance of the different variables.

Once we have done that, we are ready to get status information on the independent and dependent variables. For the dependent variables, that means measuring the subjects on the intended outcomes—and the potential side effects—that we have identified. In our example, it could mean assessing the extent to which students have equal opportunities in different areas of the curriculum.

We also need status infor-

mation on the independent variables. To what extent is teaching the same or equally good across the district? To what extent do we have qualified teachers across the district?

Next we do the most professional of all tasks in the needs assessment. We draw conclusions about which needs are being met; which ones are not being met; what opportunities within the system are being used appropriately, or misused, or not used at all; what special problems exist; and what priorities or objectives are indicated for change within the system.

Brandt: Is the needs assessment finished at that point?

Stufflebeam: No, because the audience has to take over and apply the findings. There are a number of things they might do. They might formulate a new set of goals and priorities for change in the system. They might develop an accountability system so that two years later anybody can find out what changes have taken place and why. They might also find the needs assessment results useful as they evaluate alternative program plans.

Formative and Summative Evaluation

Brandt: You put a lot of emphasis on information for decision making.

Stufflebeam: I certainly do. I think that evaluation has to be

more than an academic exercise. Unless there are important decisions to be made, we ought not to be wasting money on it. Oh, there is another reason for doing evaluations—to be accountable for what we have done. But the primary reason, in my view, is to aid decision making and thereby help us improve what we're doing.

Brandt: You distinguish between accountability and decision making as purposes. How are they different?

Stufflebeam: Well, one main difference is that evaluation for decision making is a prospective approach. The evaluator has to provide timely feedback; has to generate information the decision maker needs when the decision maker needs it.

Evaluation for accountability is a retrospective approach. There are fewer time constraints on it. We are looking back on what we did, including the decisions that were made—even the information that was available to support those decisions, and the evaluation that was done in support of the decision making.

Brandt: One way to think about evaluation is to classify it as either formative or summative. Are you saying there is a relationship between decision making and formative evaluation on the one hand, and accountability and summative evaluation on the other?

Stufflebeam: I see it that way.

Brandt: You are chairman of a committee to develop standards for evaluation. How did that come about?

Stufflebeam: Three organizations—The American Psychological Association, The American Educational Research Association, and The National Council on Measurement in Education—have had standards for educational and psychological tests since the 1950s. In the early 1970s, there was a feeling on the part of those three organizations that standards should be developed for evaluation—broadly conceived—in addition to developing new standards for tests. That concern led to formation of a planning committee and creation of a joint committee representing twelve different organizations that are now working together to create standards for judging evaluation. I have served as chairman of the committee since it began, and Phil Hosford is ASCD's representative.

Brandt: How do you expect the standards to be used?

Stufflebeam: I hope that evaluators will use them to assess what they do—their designs, their reports. I would hope that audiences—people who receive reports and depend on them—will use the standards to write better specifications and to assess reports they get.

Brandt: When will the standards become available?

Stufflebeam: The first draft is now completed and is undergoing extensive field testing, critiquing, and rewriting. We anticipate that—after all the field tests, hearings, and the revision cycles are completed—the final report will be generally available about 1981.

Brandt: I have read the draft manuscript and feel it is a major contribution. It is also very interesting reading, partly because you explain each standard by citing an illustrative case. Where did all those cases come from?

Stufflebeam: Representatives of each of the 12 organizations that are working with us nominated writers. About 40 people agreed to write for us. We gave them topics and a common format that required an illustrative case for each standard. We know that many of the cases are real. They have probably been camouflaged a bit to protect the guilty.

Brandt: All the cases are mistakes made by evaluators. Did your committee consider the possibility that publishing a book full of bad examples might give readers a poor impression of evaluation?

Stufflebeam: Yes, we considered that. We even tried to include a mixture of positive and negative cases. But we decided that showing the breach of the standard, along with guidelines for correcting the

situation, would be a better instructional approach.

Brandt: It certainly illustrates that we have had some bad evaluations. How important is it that every evaluation be itself subject to evaluation?

Stufflebeam: I think evaluators have to take their own medicine. Not only is it very important, but an outside audit can also be cost effective. Evaluators are sometimes not independent of the thing they are evaluating, so their credibility is subject to question. If they spend a modest amount to bring in someone who is independent, knowledgeable about evaluation, and credible to the audience, they add greatly to the believability of their reports.

Brandt: Looking to the future, what changes do you expect in educational evaluation?

Stufflebeam: One of the changes that is quite apparent now is that evaluation is becoming a profession. There are a number of new professional organizations developing. One is the Evaluation Network, a group of evaluators from education, medicine, social programs,

and government. Another is the Evaluation Research Society of America.

We are also going to see improvement in the practice and use of evaluation. If I didn't believe that, I would not be involved in the standards effort, because I think the standards will have some impact on that.

We are going to have better materials to learn about evaluation. We will have casebooks that do "autopsies" on completed evaluations: helping us see what was right and what was wrong, where the pitfalls were, and where the breakthroughs were in methodology.

We will probably see the emergence of more training programs in evaluation, not only for evaluation specialists, but also for teachers, administrators, and curriculum specialists. What were formerly courses in research methods in universities are being converted to evaluation training programs. There is a difference. Evaluation is a communication process, a political process; it requires management skills.

Perhaps, as the years go by, we will even see institutions coming to value evaluation so much that they budget some money to spend on it. ^[24]



Ronald Brandt is Associate Superintendent for Instruction, Lincoln Public Schools, Lincoln, Nebraska, and Chairperson, ASCD Publications Committee.

Future ASCD Annual Conferences

1978	March 4-8	San Francisco	Civic Center
1979	March 3-7	Detroit	Cobo Hall
1980	March 29-April 2	Atlanta	Georgia World Congress Center
1981	March 7-11	St. Louis	Congress Center

Copyright © 1978 by the Association for Supervision and Curriculum Development. All rights reserved.