DOCUMENT RESUME

ED 272 566 TM 860 482

AUTHOR Smith, Marshall S.
TITLE Improving the Quality and Usefulness of NCES Data.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE [Oct 85]
NOTE 8p.; In: Invited Papers: Elementary/Secondary Education Data Redesign Project, October 1985; see TM 860 450.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Agency Cooperation; *Data Collection; *Educational Assessment; *Educational Quality; Elementary Secondary Education; International Programs; Longitudinal Studies; National Surveys; *Research Design; *Research Needs
IDENTIFIERS *Educational Indicators; *National Center for Education Statistics

ABSTRACT

This paper sketches seven activities that the National Center for Education Statistics (NCES) should pay attention to over the next few years as it attempts to provide accurate and complete data on the elementary and secondary education system of the United States. The activities are: (1) act aggressively to meet the major problems of data collection and establish a system to monitor the quality of NCES data; (2) improve and coordinate the tests of academic achievement used in major surveys by paying attention to the content of the test and school curriculum; (3) the work with various parties to build indicators for assessing the equality of the educational system and the relation of the system to the nation's productivity; (4) establish an international program to compare the resources, organization and outcomes of the various advanced nations; (5) track students from the middle school years through high school and beyond; (6) consider establishing a systematic research program coordinated with the Departments of Health and Labor, the Bureau of the Census, the National Science Foundation, and the Department of Housing and Urban Development; and (7) consider carefully how to coordinate research and statistics within the Department of Education. (JAZ)

Improving the Quality and Usefulness of NCES Data

Marshall S. Smith
University of Wisconsin, Madison

Improving the Quality and Usefullness of NCES data

Marshall S. Smith
University of Wisconsin, Madison

This paper sketches seven activities that NCES should pay attention to over the next few years as it attempts to provide accurate and complete data on the elementary and secondary education system of the United States. No attempt is made in this discussion to be comprehensive and no logical or programmatic priority should be inferred from the order of items on the list.

1). Act aggressively to meet the major problems of data collection and establish a system to monitor the quality of NCES data. Over the past year there have been a number of internal (Dept. of Education) and external (GAO, "The Sorry State...") critiques of various aspects of data gathered and reported by NCES. The time is right to dedicate resources to meet these problems. One approach is to deliberately review the critics' points and make adjustments to meet each of the specific criticisms. Another approach is to work with the NAS committee reviewing NCES to examine areas that need improvement and to suggest ways of implementing corrective action. Both approaches and others should be used and a clear document spelling out what NCES intends to do and is doing about the quality of their data should be developed and released to the field for comment. Where resources are not available or where policy (legislative or administrative) gets in the way of improving the data collection thi, should be pointed out in the document and a strategy indicated for meeting the problems.

Beyond correcting the immediate problems three other steps should be taken. First, I have a hunch that there is a great deal of unnecessary and useless data collected because of legislative or administrative mandate or because it has been collected in the past and no-one has gotten around to examining its continued utility. It would be useful to ask the Academy and internal NCES staff to suggest candidates for reduction in data collection. Second, you should establish a system of yearly review. I don't have in mind anything fancy. One approach would combine internal and external review. For example, internally to NCES, on a rotating basis each year a Director of one of the NCES programs could have an assignment of producing a planning document about ways of improving data collection and data use in the agency. Or, internally to the department, each year the planning office might be requested to review a different aspect of the data collection program and use of the data in the Department. On the outside you might contract with a single particularily knowledgeable person each year to produce a short (25 page) provocative paper on one or another aspect of the data collection and analysis activities of the agency. These papers might look both at present problems and future opportunities. The effort to gain outside advice (beyond the Advisory Committee) should not end with the effort of which this paper is a part.

Third, it should be possible to establish an ongoing system for partial verification of NCES data. One approach would be to use the Quick Response survey to cross-validate certain statistics each year. Another strategy would be to work with certain key states each year to develop cross-validating systems of data collection. I do not think that you should engage in massive cross-validation. The effort should be limited and selective. Major problems signalled by discrepancies which occurred in the cross-validation would require larger efforts to correct but the frequency of major problems should be substantially reduced by such an on-going attempt to insure accuracy.

2). Improve and coordinate the tests of academic achievement used in major surveys. One of the major problems with the HSB survey is the quality of the student outcome measures, particularily the student academic achievement tests. I am mindful of the robust psychometric properties of many of the tests -- as ETS has shown in their recent report. I also understand the constraints imposed by attempting to have direct comparablity among different surveys (eg. the NLS and the HSB and presumably the NLS, HSB and HSB2) and by the need to minimize the time spent by the students taking the tests. One component of the problem rests with the lack of relationship between the content of the tests and the curriculum of schools. One version of this criticism comes from content specialists eyeballing the tests and claiming that there is little relationship between the tests and the curriculum. For example, there is little attention paid to the content of English and History--the subjects in high school in which students spend a large percentage of their time. And the tests assess only a limited conception of math and science. Another related version of the criticism comes from the observation that while the tests measure a little of what students learn in academic courses they measure nothing at all of what they learn in most general and all vocational courses. By their design, therefore, they are missing much of what high schools intend to teach. The general problem of the lack of relationship between the tests and the curriculum is manifested by the very small gains between 10th and 12th grades represented either absolutely (raw score) or relatively (percent of standard deviation of 10th grade scores).

A second problem has to do with the nature of the concepts assessed by the HSB tests. The multiple choice format, the short length of time allocated for the testing and the survey of knowledge nature of the tests reduce the chances for the measures to assess critical thinking or higher order thinking skills. Work is going on around the nation in the area of assessing higher order skills. Fredericksen's work at ETS, the ETS GRE Analytical Score efforts, and Sternberg's work for the state of Connecticut are three examples. In addition there is a lot of exploration of ways of using the computer to create testing environments that assess more than the basic skills.

A third problem is the lack of correspondence of the HSB tests with the IEA, the National Assessment or state assessments. There is, at best, scattered coordination -- the HSB with the NLS72, the National Assessment and some state assessments and, I

gather the IEA math, are examples. Someone should systematically set out the inter-relationships among the existing surveys and examine opportunities in the future so that maximum coordination (without sacrificing too much quality) could be achieved.

I recognize the difficulty in solving any one of these three problems much less all three. For one thing, NCES does not have control over either the NAEP or the IEA. My sense, however, is that there is a lot going on in the field of testing that is important, particularly in the theory and development of ways to measure complex thinking skills. There are also advances in the sampling and design areas (eg. BIB spiraling) and in the use of computers to pose complex problems and to tailor tests. Given this ferment it would be well worth NCES's time to get some people thinking about ways of ameliorating the three problems indicated above as well as other test related issues. A few papers and a couple of small conferences that include the NAEP and the IEA folks and a number of the leaders of state assessment efforts might lead to some suggestions that would vastly improve on the measures that are currently used.

3). Work with various parties to build an indicator system. This issue is already on the NCES agenda. I want to reinforce it. Over the next few years NCES should take the lead in thinking about and implementing new and better ways of assessing the health of the nation's educational system. Apart from it being part of the mandate of the agency the time is ripe. There are a number of key roles that NCES should play. NCES should be actively coordinating with the Chiefs, the National Academy of Sciences, the GAO, the NSF, the NIE testing and state and local policy centers and all of the others who now have an interest in federal, state and local level education indicators. Second, while the generic concerns of NCES are allied with those of all of the other actors (high quality data, minimize burden etc.) there are some specific federal concerns that NCES should attend to -- these have to do with indicators to assess the equality of the nation's educational system and the relationship of the system to the nation's productivity and to the nation's defense. A third area where NCES might take leadership is in the development of new strategies for developing key indicators such as ones assessing curriculum or teacher quality.

4). Establish an international program. The quality of data comparing the resources, organization, intentions and outcomes of the various advanced nations of the world is very poor. The IEA surveys, which tragically are state of the art in this area, suffer from lack of funding, proper coordination internationally and, in the US, from a lack of connection with many of the established ways of insuring adequate data collection. This combination of problems has led to erratic schedules for data collection, very poor response rates for US samples, and little use of the IEA data by the US policy system. One way of beginning to ameliorate these problems would be to have NCES assume responsiblity for coordinating US involvement in IEA activities.

Other than the IEA there are only scattered sets of

international comparisons which involve the US. Some are carefully carried out though they have only limited generalizability because of sample breadth (for example, Harold Stevenson's work) -- others suffer from major problems (the Texas newspaper's study). I don't believe that NCES should have much to do with these entrepreneurial efforts except to keep track of them.

There might also be opportunities for cooperation with OECD or with other nations, singly or in groups. My instinct is that there ought to be a mechanism within NCES to seek out these chances -- this should not be institutionalized in a position, however. The Director of NIE used to sit on the CERI-OECD board. This might be an avenue for coordination.

5). Track students from the middle school years through high school and beyond. Analyses of data from HSB have demonstrated the importance of tracking students while they are in high school to document the effects of schools on student achievement and attitudes. Even if the results of cross-sectional analyses on 12th grade outcomes are corroborated by longitudinal analyses of 10th to 12th grade gains the added confidence in the results that comes though the longitudinal addition of earlier grade test scores and attitudes is worth the investment. And, occasionally, it turns out that longitudinal studies provide a very different set of findings than did cross-sectional data. Beyond that, longitudinal data collected in 10th through 12th grade in HSB, in contrast to the data collected only in 12th grade and beyond in the NLS, allowed researchers to begin to examine nationally such issues as students' reasons for dropping out and the short and long term effects of dropping out.

Unfortunately, 10th grade is not early enough to fully examine the effects of high schools on student achievement and attitudes or the reasons of students for and consequences of dropping out, among other issues. By 10th grade many students, particularily poor students in inner cities, have already either dropped out officially or stopped attending school. According to HSB data, for example, 16% of the nation's students drop out before graduation from high school -- most other national estimates which include dropping out prior to 10th grade place the percentage of dropouts as 25-28%. The exclusion of upwards of 40% of the nation's dropouts from the HSB analyses and, in particular, those dropouts who left school prior to 10th grade, suggests that profiles, analyses and policy conclusions developed from the existing HSB data may be misleading if they are interpreted as representing the entire population of dropouts.

Our ability to analyze and understand other high school processes is also limited by studies which gather initial data on 10th graders. The phenonema of tracking and ability grouping is well underway prior to 10th grade in almost all high schools--the determinants of these assignment practices are operating by 7th and 8th grade. The same arguement applies to the comparison of public and private high schools--most of the selection issues have been settled long ago by the time students reach the 10th grade. Still another area of study that is limited by studies which begin in 10th grade have to do with determinants of

patterns of participation in extra-curricular activities and of work outside of school. The general point is that inferences about the effects of high schools on students are necessarily limited if analysts do not have data on students that preceeds the entrance of the students into high school.

A related substantive arguement has to do with our collective lack of knowledge about the experiences of students in the middle school years -- grades 6th through 9th. With the exception of the IEA studies(which are very erratic) and the 13 year old samples from NAEP (also erratic, cross-sectional and only recently containing much school and teacher data) we have no systematic data on the experiences, quality of schooling and achievements of middle level school students. If the new high school longitudinal survey were started in the 7th or 8th grade we would gain considerable data about the middle school years, the transition between middle and high school and get a more complete picture of the high school than we had previously.

There are substantial issues having to do with cost, time, and comparability of data sets which would have to be resolved for a program of research like this to be undertaken. Major among these are the various routes that students take from middle level to high schools and the costs of tracking students. My sense is that there are a variety of strategies that could be developed to gain information about efficient ways of collecting middle to high school data -- maybe a small working conference would be useful to do some preliminary planning on this issue.

6). Consider establishing a systematic research program coordinated with Health, Labor, the Census, NSF and HUD. This proposal stems from two motives. First, I am sure that there are inter-governmental coordination committees that meet every two months to share information to help all interested parties in the government stay abreast of the latest and most effective methods of data gathering and analysis. I am equally sure that although these committees are sometimes very effective that a lot more coordination could be done. Second, I am confident that there are a substantial number of ideas in the field about ways of improving data gathering and analysis strategies that could use some stimulation and direction and could provide great savings to the federal government in the very near future. On the data analysis front I have in mind as examples the matching strategies recently examined by Rubin and others and the modeling estimation techniques proposed by Heckman. On the sampling and data collection side the BIB spiraling approach developed by ETS to collect NAEP data is a recent example. Strategies for collecting indicators represents another area that requires systematic research.

In my experience one of the best ways to get agencies to share expertise is to involve them in a collective activity. Of course the activity has to be in an area of concern. And it has to require some committment of resources from each agency. A joint research program among the primary data gathering offices of each of five or six agencies such as Labor, the Census, HUD, the Health statistics office and NSF could be carried out with a modest contribution (say $200K) from each agency for each year.

This would put $1 million in the field to directly address problems that would make the efforts of the agencies more effective and efficient. I suspect it would also lead to greater communication among agencies in data collection and analysis strategies. It might also lead to a greater understanding among the various agencies about the nature and content of data collected by other parts of the government.

7). Consider carefully how to coordinate research and statistics within the Department of Education. Over the 13 years of NIE's existence there has been little systematic cooperation between it and NCES. In the past six months this may have changed. At first blush my sense is that greater cooperation between the agencies and, indeed, systematic coordination of research funding and the gathering and analysis of statistics would be of great benefit and little cost. Such coordination might lead to combining the functions under a single Director. As far as I know, however, the research and statistics gathering and analysis functions are separated in other departments of the federal government. This may be due to the fact that the constituent communities are different, each wanting their own small agency. Or it might be due to a desire to separate research from the on-going policy orientation that the statistics agencies often have. Neither of these reasons appear compelling to me but there may be other reasons to keep research and statistics at arms length from each other. With the NIE being reauthorized and organizational changes to OERI being proposed internally it seems to be a proper time to think through this issue.

END