DOCUMENT RESUME

ED 131 087 95 TM 005 670

AUTHOR Coulson, John E.
TITLE Attrition Effects in Large-Scale Longitudinal Studies.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE [Apr 76]
CONTRACT 300-75-0332; OEC-0-73-0831; OEC-0-73-6336
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Anti Segregation Programs; Compensatory Education Programs; Dropouts; *Longitudinal Studies; Program Evaluation; Research Design; Research Problems; *Statistical Analysis; *Statistical Bias; *Testing Problems
IDENTIFIERS *Attrition (Research Studies)

ABSTRACT

Using examples from evaluations of the Emergency School Aid Act (ESAA) Basic Grants Program, the ESAA Pilot Program, and the sustaining effects of compensatory education programs, school and student attrition are discussed. As in the example cases, appreciable attrition can be expected in most longitudinal studies. The possible effects of this attrition on descriptive analyses, analyses of student gains for each school year, and analyses of differential achievement gains for different treatment groups are considered. The decision of whether to use statistical adjustments for bias is also analyzed. (BW)

ATTRITION EFFECTS IN LARGE-SCALE LONGITUDINAL STUDIES

John E. Coulson

System Development Corporation

The work upon which this publication is partially based was performed pursuant to Contracts OEC-0-73-0831, OEC-0-73-6336, and 300-75-0332, with the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the positions or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

Prepared for presentation at the annual meeting of the American Educational Research Association, San Francisco, April 20, 1976.

ATTRITION EFFECTS IN LARGE-SCALE LONGITUDINAL STUDIES

John E. Coulson

System Development Corporation

Before discussing some of our actual problems and approaches in dealing with attrition in longitudinal studies, I would like to describe briefly the two studies that will serve as the focus of discussion in this paper. The first is an evaluation of the Emergency School Aid Act (ESAA), which System Development Corporation is conducting under a contract from the U.S. Office of Education. Under the ESAA Basic Grants Program, Federal funds are provided to eligible school districts to help them meet special needs incident to desegregation. Schools receiving Basic program funds have multi-racial enrollments. Activities conducted with these funds include both desegregation-related projects, such as training of staff members in race relations, and efforts to enhance students' academic skills. The ESAA Pilot Program, by contrast, is directed toward schools with 50% or more minority enrollments; often, in fact, schools receiving Pilot program funds are 90% to 100% minority. In Pilot award schools the ESAA funds are used entirely for compensatory education purposes, particularly the improvement of students' reading and mathematics skills.

The ESAA Evaluation Project, which the present author directs, attempts to assess the impact of ESAA funding on student achievement, on students' perceptions of discrimination in the schools, and on the degree of minority group isolation in the funded districts. For this purpose the study includes a true experimental design, with pairs of ESAA-eligible schools selected in each sample district, and with the schools in each pair randomly assigned to treatment (ESAA-funded) and control (non-ESAA) conditions. At each school, sample students were randomly drawn in the initial year from across classes, in the third, fourth, and fifth grades of elementary schools, and the tenth, eleventh, and twelfth grades of secondary schools. In each subsequent year,

Prepared for presentation at the annual meeting of the American Educational Research Association, San Francisco, April 1976.

testing was conducted at those same grade levels, dropping students who had moved to the sixth and twelfth grades, and drawing new samples at the third and tenth grade levels. Within a grade level, students continuing in the longitudinal sample were augmented by replacement students in the second year to keep the cross-sectional samples at each school and grade sufficiently large. Achievement tests and school climate questionnaires are administered at the beginning and end of each year in treatment and control schools.

The ESAA evaluation also seeks to identify critical factors in the local school projects that contribute to or interfere with program success. Each year, questionnaires are administered in both treatment and control schools to collect data on program, staff, and student characteristics. The ESAA evaluation contract is now in its third and final year of data collection. A Year One Report has been issued, and the Year Two Report is presently in preparation.

The second study used for illustrative purposes in this paper is a multi-year evaluation of the sustaining effects of compensatory education programs, including but not limited to the E.S.E.A. Title I Program. This study is being conducted by SDC under the direction of Dr. Launor Carter; like the ESAA study, it is funded by USOE. The Sustaining Effects Evaluation is currently in its planning and design phase, with the first of three years of data collection scheduled to begin this fall. Standardized achievement tests will be administered at the beginning and end of each year to all students in over 200 schools. Questionnaires will be used to collect data on district, school, program, staff, and student characteristics. There will be no true control group--that is, no random assignment of schools or students to treatment and control conditions--but efforts will be made to use natural variations in poverty level and in funding level and type of compensatory education.

A. SCHOOL AND STUDENT ATTRITION IN ESAA AND SUSTAINING EFFECTS STUDIES

1. Attrition of Schools

Given this background on the ESAA and Sustaining Effects evaluations, it may be of interest to consider the types of attrition problems that we anticipated and, in the case of the ESAA study, to compare those with problems actually encountered. First, a certain loss of whole schools from the sample was

expected in both studies, although it was difficult to predict the magnitude and consequences of this loss. In the ESAA study, emphasis was placed on finding districts with pairs of similar, ESAA-eligible schools, so that members of each pair could be randomly assigned to the treatment and control groups. It turned out to be quite difficult to find enough suitable pairs of schools for the Pilot elementary and the Basic secondary subsamples. With respect to school attrition, therefore, we were particularly concerned about how such attrition might affect our longitudinal impact analyses for these subsamples. We felt that school-level attrition might occur either because of certain districts having their ESAA funds discontinued, or because of refusal or inability of some schools to continue participation in the study.

In the actual event, school-level attrition in the ESAA study was less serious than expected for cross-sectional impact analyses but quite a problem for longitudinal analyses. We lost 19 of our 78 original pairs of schools by the second year of data collection, in almost all cases because ESAA funding in the districts involved had been discontinued. However, by drawing on newly funded projects we were able to bring our total second-year samples back to 20 Pilot elementary pairs, 34 Basic elementary pairs, and 20 Basic secondary pairs. At the same time, unfortunately, our two-year longitudinal samples dropped down to 33 Basic elementary pairs, 9 Basic secondary pairs, and 12 Pilot elementary pairs. Thus, the Basic secondary and Pilot elementary samples have fallen to levels that make longitudinal impact analyses for those subsamples of somewhat reduced value.

In the new study of Sustaining Effects of Compensatory Education, we anticipate fewer problems of school-level attrition. In part, this is because of greater year-to-year stability in the funding of Title I and similar compensatory education programs, as compared with ESAA funding. Also, because of the relatively large number of sample schools in the Sustaining Effects study, there seems less danger that the loss of a few schools will create serious imbalances in the sample. We are taking a number of steps to try to reduce

any tendency of districts or schools to drop from the study. First, we are providing special incentive payments to school administrators and staff who participate by administering tests or completing questionnaires. These payments, while not equivalent to the school's regular hourly wages, are, nevertheless, fairly substantial, amounting to several hundreds of dollars for some participants. Second, we have tried to keep questionnaires as brief as possible to reduce respondent burden, as we have found this a key factor in the school staffs' feelings toward an evaluation. Another procedure that we have found helpful in our ESAA evaluation and plan to continue in the Sustaining Effects study, is to invite a key administrator from each sample district, usually the local program coordinator, to Santa Monica once or twice each year for orientation sessions. These meetings not only serve the orientation function, but give the local program coordinators a chance to get away from their normal program chores and to enjoy a not-too-stressful day in Southern California at no cost to the districts.

2. Attrition of Students

Now what about attrition at the student level? At the time we were starting the ESAA evaluation we did a fair amount of literature research and spoke to a number of people we believed would be knowledgeable, trying to get some realistic estimates of how great our attrition problem might be. We found a great deal of folklore, a strong mystique, and very little hard, factual information to help us. One problem was that most previous longitudinal studies have tested only in the spring of each year, so there has often been no clear distinction between attrition occurring during a school year, and that occurring over the summer months. School district records were more often confusing than helpful, because of the districts' preference for speaking of "transiency" rather than of "attrition." Since transiency rates take account of each student entering the school system as well as each student leaving, they tend to give artificially inflated pictures of the amount of instability in district and school enrollments.

What was lacking in empirical data was more than made up in totally speculative but uniformly gloomy predictions about student attrition. The ESAA sample is heavily weighted by large inner-city schools with large percentages of minority and disadvantaged students. In such schools, we were consistently

told, student turnover is rampant. We could expect, according to most soothsayers, to lose at least one third to one half of our sample each year.

In the face of these dire predictions we took several steps designed to keep our student samples as large and as stable as possible within the project's funding constraints. Makeup sessions were given for the post-test administrations of the achievement test and school climate questionnaires. We did not depend entirely on mail to obtain questionnaire responses, but sent field representatives into each sample school to encourage and aid the respondents, including students, in completing their questionnaires. We also increased the number of third and tenth grade students in the first-year sample from 30 to 60 students per school, so that sample sizes for those groups would still be adequate at the end of third evaluation year. Finally, we instructed all of our data collectors to maintain careful logs on each sample student, so that it was immediately obvious when we had excessive attrition in any particular classroom or school. This enabled us to take quick remedial action where needed, such as sending a backup data collection team to the school in question to locate and test some of the "missing" students. These special visits were above and beyond the regularly scheduled makeup tests administered routinely as part of the basic testing procedure.

One factor considerably confounding the whole issue of student attrition has been the recent legislation requiring protection of student data. To meet this requirement, in the second year of ESAA data collection, we introduced anonymity procedures designed to guarantee that neither SDC nor USOE would be able to associate student names with any student response data, either on tests or on student questionnaires. Under these procedures, SDC can associate student ID numbers with student response data, but never receives any information about student names. The schools, in turn, maintain records associating student names and ID numbers, but are never given individual student response data. A trust company in St. Louis serves as escrow agent, maintaining a duplicate copy of the lists associating student names and ID numbers, in case a school loses its master list or in case there is some dispute as to whether the "right" student has taken a test.

These anonymity procedures appear to have worked quite successfully, but they do introduce delay and some uncertainty into the picture whenever efforts are made to "recapture" lost students. Rather than simply calling the schools and saying that we would like another chance at testing Jimmy Jones and Robert Brown, we must read ID numbers to someone in the district who then translates those numbers into names. When our test supervisor arrives back at the school he must again deal in ID numbers, and must take on faith that students appearing before him correspond appropriately to the ID numbers given. On the positive side, however, the protection of respondents' identities--and this applies to school personnel as well as students--may have led them to respond more freely and honestly to questionnaire items asking about personal and demographic characteristics.

Despite various problems, student attrition in the first ESAA evaluation year was gratifyingly low--much lower, in fact, than we had been led to hope for. For the overall sample, around 90% of the students pretested in fall 1973 were also posttested in the spring of 1974. Since we did not take any extreme measures to track down all absent students, it seems quite certain that the percentages of students who had actually left the schools were even smaller than indicated by our test records.

While attrition over the first school year was encouragingly low, the loss of students over the summer months was another story altogether. Roughly two-thirds of the elementary students pretested in the first year of the evaluation were pretested in Year Two. Assuming that attrition during the second school year was similar to that for the first year, as our preliminary analyses suggest, approximately four out of every ten students pretested in Year One were lost from the sample by the Year Two posttest. These loss figures do not take into account the cases where entire schools dropped from the study, but are based on the set of schools participating in both the first two years of the evaluation.

The Sustaining Effects Study, in contrast to the ESAA evaluation, will draw more broadly from the total spectrum of schools across the country, and will thus include some fairly affluent schools as well as schools with large

proportions of disadvantaged students. Overall, therefore, we do not expect as high an attrition rate in the Sustaining Effects study as we have found in ESAA schools. Nevertheless, there is a distinct possibility that differential student attrition rates in poverty schools and affluent schools may, over the three years of data collection, create a progressive change in the composition of the longitudinal sample.

B. POSSIBLE EFFECTS OF ATTRITION ON DIFFERENT TYPES OF ANALYSES

Given that some appreciable attrition can be expected in most longitudinal studies, it is extremely important to identify the types and magnitude of bias that may be introduced by that attrition in a particular evaluation setting. Only when the bias is clearly defined can reasonable judgments be made concerning strategies for dealing with its effects, yet many studies fail to report any detailed data on bias or to discuss possible implications of that bias for interpretation of findings.

Going one step further, attrition effects and resulting bias should be considered in the context of specific evaluation goals and specific analytic models. Attrition may not introduce substantial bias, or it may produce different kinds of bias that require different methods of handling depending on the types of analyses that they influence. Developing this point in any detail is beyond the scope of this paper but, for illustrative purposes, consider briefly three kinds of analyses called for in both the ESAA evaluation and the Sustaining Effects study. (There are other kinds being performed, but for simplicity this paper's discussion is limited to these three.) First, there are descriptive summaries, consisting of frequency distribution, means, standard deviation, and similar statistics for a variety of school, program, staff, and student characteristics. Second, there are analyses of student achievement gains for each school year, and cumulatively across years. And third, there are impact analyses, or analyses of differential achievement gains for different treatment and/or control groups.

Looking first at descriptive analyses, sample attrition may reduce the degree to which descriptive statistics are representative of some population or universe. Adjustments for bias in descriptive statistics can be quite straightforward to make, by means of a weighting process designed to estimate

or recreate the relevant properties of some hypothetical sample composition, such as a nationally representative sample. Such a procedure presupposes, however, that distributional properties of certain key variables used for weighting purposes are known for the referrent universe as well as for the sample in question, and such data may not automatically be available.

Next, consider bias in the context of analyses of achievement gains. A sample may be substantially biased from the standpoint of certain descriptive statistics, without the bias necessarily having harmful effects on the validity of statistics characterizing students' achievement gains or program impact. The key question, of course, is whether the descriptive characteristics in which bias has been found are associated with achievement level or achievement gains.

Finally, taking this logic one step further, sample bias, whether inherent in the original sample design or resulting from attrition, may distort our picture of overall student achievement gains without necessarily invalidating analyses of relative gains in different comparison groups. Again, the question is whether bias has been introduced in variables differentially influencing the gains of those comparison groups.

C. BIAS IN THE ESAA EVALUATION

For a more concrete and less hypothetical view of bias problems resulting from attrition, a further look at the ESAA evaluation data may be useful. We have not yet completed our analyses of the two-year longitudinal data, but do have a fairly clear picture of bias resulting from attrition within the first evaluation year. Briefly stated, we found that students lost to the sample between the pretest and the posttest included disproportionately high percentages of disadvantaged students, minority students, and students with low pretest scores. It therefore appears that our analyses of student achievement scores, based on students receiving both pretest and posttest, may systematically overestimate the performance of students in the sample schools. Furthermore, regressions of posttest scores on pretest scores indicate that students with lower pretest scores also made smaller achievement gains over the school year. Thus there is reason to presume a systematic

overestimation in our analyses of achievement gains for students in the sample schools. By contrast, we did not find evidence of bias in our analyses of relative achievement gains in treatment and control groups. This last point may be somewhat academic, however, since we also found no evidence of overall differences in achievement gains between the two comparison groups.

D. CONSIDERATIONS IN DECIDING WHETHER TO USE STATISTICAL ADJUSTMENTS FOR BIAS

The ESAA evaluation's use of tests at both the beginning and end of each school year yields extremely valuable information about the amount and nature of bias introduced by sample attrition. When only posttests are used, as is true in many evaluations, bias in the students' entry scores can only be indirectly inferred from information about bias in other student characteristics such as race or socioeconomic status. But even when bias can be clearly identified and quantified, adjusting for that bias is by no means simple or straightforward. Such adjustments require important assumptions to be made about the biased variables and about their interactions with other variables that may or may not themselves be biased. One such assumption is that a linear relationship (or other known relationship that can be fairly precisely specified) exists between the variable being adjusted and the variable on which adjustments are based. Another is that one can extrapolate or interpolate from the observed behavior of students remaining in the sample to the behavior that would have been exhibited by the students who dropped out of the sample. Often it is difficult to determine how fully the assumptions are met, and thus to estimate the error that may be introduced through the adjustment procedures. In the ESAA study, adjustments for bias could have been made by using pretest scores as a covariate in variance analyses of the pretest-posttest gains. However, such adjustments almost inevitably introduce their own errors, especially when the students missing from the posttest are predominantly from one extreme of the pretest distribution. Since the actual attrition rate in the first year was so low, we felt it wiser not to attempt to adjust for bias, but simply to describe the bias and warn the reader of its probable effects on our analyses of gains and program impact. As the attrition problem grows in magnitude over the second and particularly the third year of data collection and analysis, we may

decide to make adjustments for bias as the lesser of two evils. The important point to recognize is that decisions involving possible adjustments for bias do represent trade-offs, and one does not automatically gain by performing such adjustments. I believe, in fact, that next year's AERA meeting could profitably include a technically oriented symposium focusing specifically on factors that may argue for or against use of alternative bias adjustment procedures under particular research and evaluation conditions.

One other issue, which may be illustrated by the Sustaining Effects study, relates to the possible use of replacement students in synthetic or quasi-longitudinal analyses. In the Sustaining Effects evaluation, all students present in the sample schools are tested at the beginning and end of each school year. A synthetic cohort might be created by using replacements as well as continuing students to calculate pretest-posttest achievement gains. This approach might not only help to maintain sample size, but might also help to reduce bias due to sample attrition. But to what extent will such analyses give an accurate picture of gains that would have resulted had the original sample been maintained intact and without replacements? The answer to this question seems to depend on several factors, of which two are particularly critical. First, how similar are the replacement students to those lost to the study through attrition? There is some evidence from previous studies, including the ESAA evaluation, that the replacements are likely to closely resemble the attritted students in terms of distributions on socioeconomic status, race, and other background variables. (The exception might be schools in communities experiencing very rapid changes in racial or socioeconomic composition.) In general, then, we expect the processes of attrition and replacement in the Sustaining Effects study to maintain a sort of "steady state" with respect to key student characteristics. We are less confident, however, about a second issue affecting the legitimacy of synthetic longitudinal analyses. This issue relates to the shape of students' growth curve over progressively longer exposure to the program or programs in question. The use of synthetic longitudinal analyses is maximally defensible when achievement gains are linearly related to exposure interval--that is, when any given period of program participation has the same effects on achievement as any other period. The synthetic cohort approach does <u>not</u> seem valid in

cases where the achievement/exposure slope is significantly curvilinear. Suppose, for example, that a program produces rapid achievement gains over the first few months of a student's participation, but then the achievement tends to flatten out. In such a case, the use of replacement students in synthetic longitudinal analyses would tend to overestimate the program's impact over a two- or three-year period. Conversely, in the case of a positively accelerating achievement/exposure curve, the use of replacement students might underestimate program impact. In the Sustaining Effects study, growth curves of the true longitudinal samples will be examined to determine the degree of credence that can be placed in synthetic longitudinal analyses. If the results appear encouraging, longitudinal analysis of student performance will be performed both with and without the replacement students. If there are differences in the findings for these two sets of analyses, attempts will be made to reconcile and explain these differences.