DOCUMENT RESUME

ED 271 508 TM 860 419

AUTHOR Hilton, Thomas L.; And Others
TITLE Score Change Partitioning Analysis of NLS and HSB Test Data.
PUB DATE Oct 84
NOTE 11p.; Paper presented at the Annual Meeting of the Evaluation Network/Evaluation Research Society (San Francisco, CA, October 10-13, 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Covariance; Cohort Analysis; Cross Sectional Studies; Educational Trends; Family Environment; High Schools; Institutional Characteristics; *Longitudinal Studies; *Path Analysis; Predictor Variables; Research Methodology; School Demography; *Scoring Formulas; Statistical Studies; Student Attitudes; Student Behavior; *Trend Analysis; Weighted Scores
IDENTIFIERS *High School and Beyond (NCES); *National Longitudinal Study High School Class 1972

ABSTRACT

Since the mean score for a sample composed of several subgroups can be viewed as the sum of the mean of each subgroup weighted by the proportional size of the subgroup, then the mean change in a time period--in this case, from 1972 to 1980--is the sum of the differences between the means for each subgroup, with each mean weighted by its proportional size. Thus, the mean change for the total sample is influenced both by mean change in the score for each group--in this case, a test score--and by change in the proportional size of the group. A decline or gain in mean test scores has both a subgroup mean component and a population shift component. In addition, how much a particular subgroup may contribute to the total change can be shown to depend on what might be called its weighted distance from the original total mean. How these simple concepts can be used to decompose score gains or declines is demonstrated by means of data from the seniors in the National Longitudinal Study (NLS) High School Class of 1972 and High School and Beyond (HSB) 1980. The presentation should be of both methodological and substantive interest. (Author/LMO)

Score Change Partitioning Analysis of NLS and HSB Test Data

Thomas L. Hilton, Judith Pollack, Donald A. Rock

Educational Testing Service

Since the mean score for a sample composed of several subgroups can be viewed as the sum of the mean of each subgroup weighted by the proportional size of the subgroup, then the mean change in a time period--in our case, from 1972 to 1980--is the sum of the differences between the means for each subgroup, with each mean weighted by its proportional size. Thus, the mean change for the total sample is influenced both by mean change in the score for each group--in our case, a test score--and by change in the proportional size of the group. A decline or gain in mean test scores has both a subgroup mean component and a population shift component. In addition, how much a particular subgroup may contribute to the total change can be shown to depend on what might be called its weighted distance from the original total mean. How these simple concepts can be used to decompose score gains or declines was demonstrated by means of data from the seniors in NLS 72 and HSB 80. The presentation should be of both methodological and substantive interest.

Score Change Partitioning Analysis of NLS and HSB Test Data

Thomas L. Hilton, Judith Pollack, Donald A. Rock

Educational Testing Service

To draw conclusions about why individuals change over time requires a tightly controlled experiment wherein all the factors that are not of interest are held constant and only the hypothesized causes are allowed to change. Longitudinal studies, wherein the variables not of interest are held constant by statistical means, are a second-best substitute for controlled experiments, for at least the same individuals are measured at the beginning and at the end of the data collection. But drawing conclusions about causes of individual change from cross-sectional data is impossible in any rigorous sense. Nevertheless, when one is confronted with questions of pressing national importance and only cross-sectional data are available, one does the best one can. In the present study this consisted of three approaches: (1) examination of changes in mean test scores and partitioning of the mean change into components of interest, (2) analysis of covariance partitioning, and (3) path analysis.

The first technique describes the extent of the relationship between selected population and behavioral shifts to score decline. This type of analysis provides considerable detail about how classifying an individual on one or two variables at a time relates to test score changes between 1972 and 1980. In addition, this methodology attempts to partition the total score change into that part that was due to population shifts in the classification variables and that part that was due to mean changes within the classification groupings.

This partitioning procedure, however, does not lend itself to evaluating the impact of any one given classification variable or set of classification variables on score changes while controlling for the effects of numerous confounding variables. The second procedure that was used attempts to look at the relative impact of selected blocks of variables on the 1972-1980 mean score changes both before and after controlling for other confounding blocks of variables. The four blocks of variables are (1) demographics (e.g., race/ethnicity, sex); (2) student behavior and attitudes; (3) school characteristics; and (4) home educational support systems (e.g., parental influences, parental education, etc.). This partitioning procedure uses a "step down" analysis of variables that form a block while controlling for the remaining blocks. A second step in this step down ANCOVA is the identification of the variables in each block that contribute the most to that particular block's net effect on mean score change.

In a sense this method takes the multiplicity of findings from the first method and sorts them into logical sets or blocks and summarizes their net impact on mean test score change.

The above two methods are primarily exploratory and descriptive in nature. The third and final method contrasts path analysis models for the 1972 and 1980 cohorts separately in an effort to shed light on what changes in processes might have occurred to account for both the overall and, as well, differential score decline.

<u>Partitioning of mean score changes</u>. It is well known that the overall mean for two or more subgroups can be viewed as the sum of the subgroup means when each subgroup mean is weighted in accordance with its proportion of the total group. In other words,

$$\bar{X}_T = (p\bar{X})_1 + (p\bar{X})_2 + \cdots\cdots$$

Where $\bar{X}_T$ is mean for total sample and p is proportional size of each subsample.

Furthermore, a decline or gain in some time period--in the present case, from 1972 to 1980--is as follows:

$$\bar{X}_{T80} - \bar{X}_{T72} = (p_{80}\bar{X}_{80} - p_{72}\bar{X}_{72})_1 + (p_{80}\bar{X}_{80} - p_{72}\bar{X}_{72})_2 + \cdots\cdots$$

Thus, the total mean can change as a result of either a change in the mean of a subgroup or a change in the proportional representation of that subgroup. We will refer to the first component as the group mean change or <u>G</u> and the second as the population shift change or <u>P</u>. To estimate the magnitude of these components, two different calculations are made. The first component is calculated by applying, for each subgroup, the group's proportion in 1972 to the group's mean in the 1980 population and then summing over all subgroups. The result is the mean score in 1980 that would have been expected if each subgroup's representation in the population had not changed but its mean had changed. The difference between this number and the observed 1980 mean is the "Change due to subgroup mean changes" or <u>G</u>.

Similarly, the "Change due to population shift" or P is calculated by applying, for each subgroup, the group's proportion in the 1980 population to the group's mean score in 1972, and then summing over all subgroups and proceeding as with G.

The sum of P and G is not equal to the total mean change for the grouping variable, for there may be an interaction between the two components. This interaction component (I) is calculated simply by subtracting the sum of P and G from the total mean change. In most cases, the term is a negligibly small number, and even where it is not its meaning is so difficult to interpret that the authors have not attempted to do so.

Moment of group mean change. When, in a given time period, the means for two or more subgroups change and, in addition, each subgroup's proportion in the population changes, it is not obvious how much each subgroup may have contributed to the total mean change. For this reason, a fourth statistic was computed that is referred to as the moment of the subgroup mean (M) because of its similarity to the familiar moment or torque in physics. This is the difference between the group's "weighted distance" in 1980 and in 1972. This distance is computed as the product of the group's proportional representation in its own year times the deviation of the group's mean from the 1972 grand mean. The moment, M, or, more precisely, the change in moment may be expressed as

$$\Delta M = p_2(\overline{X}_2 - \overline{X}_{1T}) - p_1(\overline{X}_1 - \overline{X}_{1T})$$

where p_2 and p_1 are the proportions at time 2 and time 1, $\overline{X}_{1T}$

is the total mean at time 1, and $\bar{X}_a$ and $\bar{X}_1$ are the subgroup mean at time 2 and time 1. The sum of the Δ M's or the change in the weighted distances over all groups is the distance of the whole population from the original mean, i.e., the difference between the 1980 and 1972 grand mean. A large number in the "Partition due to group" column can be a result of a population change for a highly deviant group (an increased proportion for a low-scoring group or decreased representation of a high-scoring group) and/or a score decline for a group that is unusually large relative to other groups. Interpretation of the score partitions must be in conjunction with score mean and proportion statistics, as well as the "Change due to population shift" index.

In discussing the results for a particular grouping variable, e.g., race/ethnicity, we frequently will refer to the contribution of a particular category or subgroup, e.g., White students, to the overall decline, judging from the change in the subgroup's moment. It should be understood in all cases that the apparent causation may well be spurious; that is, the change in the variable in question may be a reflection of a change in some more fundamentally causative variable with which it is correlated. In the interest of brevity we usually will omit the proper qualifications.

<u>Some comments on mean score change</u>. To assist the reader in interpreting mean score changes for a group from 1972 to 1980, the following comments are offered. The mean score for a particular subgroup of the total sample can change from 1972 to 1980 for at least four possible reasons:

1. The members of the category may in 1980 represent a smaller more selected sample of the population than in 1972. For example, if the dropout rate for a category of students was much higher in 1980 than in 1972, then it is reasonable to expect either less score decline or even a score gain from 1972 to 1980 if the students dropping out represented a lower ability stratum of the population, as usually is the case. A substantial decrease in the size of the subgroup from 1972 to 1980 relative to other subgroups would be a warning that the 1980 subgroup may have been influenced by differential attrition for the period.

2. A difference in the mean score for a subgroup may reflect a change in the personal characteristics of the population of students represented by the subgroup category. For example, in 1972 the category "Suburban students in academic programs" may have been predominantly White students, whereas in 1980 the category may have included a significant fraction of non-White students.

3. From 1972 to 1980 the classroom, school, or community environment may have changed in such a way as to change the behavior of the students at the two times.

4. The meaning of the category may have changed. For example, in 1972 the category "Less than 5 hours of homework" may have meant that the average category member did 4 hours of homework each week whereas in 1980 the members of that category did an average of 2 hours of homework each week. In other words, the authors may have been only partially successful in holding certain independent variables constant by subdividing the total group into categories. Each of these possible interpretations are exemplified in the results reported in the next section.

Table 1

Mean IRT Vocabulary Scores in 1972 and 1980 by Sex, with Partitioning

	Year		1980-1972
	1972	1980	Difference
A. Mean Vocabulary Score			
Male	6.44	5.90	-.54
Female	6.67	5.69	-.98
Total	6.55	5.79	-.77
B. Weighted N			
Male	1,425,843	1,218,450	-207,393
Female	1,433,577	1,355,969	- 77,608
Total	2,859,421	2,574,419	-285,002
C. Proportion of High School Senior Population			
Male	.50	.47	-.03
Female	.50	.53	.03
D. Change Due to Population Shift (P)		0.01	
E. Expected Change with No Population Shift (G)		-0.76	
F. Total Change Due to Interaction (I)		-0.01	
G. Partition of Total Change Due to Group			
Male		33%	
Female		67%	

Table 2.1

Differences Between 72-80 Senior Test Score Means When Each Block is Separately Held Constant

	VOCABULARY		READING		MATH	
OBSERVED MEAN DECLINE	.85		1.04		1.04	
MEAN DECLINE OR INCREASE AFTER ADJUSTING FOR:						
		NET CHANGE		NET CHANGE		NET CHANGE
DEMOGRAPHICS ONLY ...	.63	.22	.79	.25	.60	.44
STUDENT BEHAVIORS ONLY...	.28	.57	.48	.56	.56	.48
SCHOOL CHARACTERISTICS ONLY ...	.21	.64	.24	.80	.23	.81
HOME SUPPORT ONLY...	1.06	-.21	1.31	-.27	1.45	-.41

TABLE 2.2

ADJUSTED MEAN DIFFERENCES FOR 1972-1980 TEST SCORES BY SELECTED BLOCKS OF EXPLANATORY VARIABLES