DOCUMENT RESUME

ED 063 881 HE 003 133

AUTHOR	Boruch, Robert F.; Creager, John A.
TITLE	Measurement Error in Social and Educational Survey Research.
INSTITUTION	American Council on Education, Washington, D.C. Office of Research.
REPORT NO	ACE-RR-v7-2
PUB DATE	May 72
NOTE	66p.
AVAILABLE FROM	Office of Research, American Council on Education, 1 Dupont Circle, Washington, D. C. 20036
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Educational Research; *Higher Education; Predictive Validity; *Research Methodology; *Statistical Analysis; *Validity

ABSTRACT

The Cooperative Institutional Research Program (CIRP) of the American Council on Education is an example of a systematic program of social and educational research in which survey methods using questionnaires constitute a major technique for collecting data. This longitudinal research program is dedicated to the study of higher education, particularly its impact on student development. In such a program, measurement error may affect data analysis and interpretation at each state of the research system through which data are transmitted and processed. That is, inaccuracy or imprecision in recorded response and in the statistics based on these data may occur not only during the solicitation of data but also during information transcription and file development. The purpose of this report is to present empirical results about certain problems of measurement error in an educational survey system. These results are confined to measurements derived from items in the annual surveys of entering freshmen that are a basic part of the CIRP. After discussion the meaning and consequences of measurement error, the paper presents a review of the pertinent literature, and then presents the results of recent empirical investigations. (Author/HS)

ED 063881

OFFICE OF RESEARCH · AMERICAN COUNCIL ON EDUCATION

Measurement Error In Social and Educational Survey Research

ROBERT F. BORUCH
JOHN A. CREAGER

VOL. 7 NO. 2 · 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

HE003133

AMERICAN COUNCIL ON EDUCATION

Roger W. Heyns, *President*

The American Council on Education, founded in 1918, is a council of educational organizations and institutions. Its purpose is to advance education and educational methods through comprehensive voluntary and cooperative action on the part of American educational associations, organizations, and institutions.

The Council's Office of Research was established in 1965 to assume responsibility for conducting research on questions of general concern to higher education. *ACE Research Reports* are designed to expedite communication of the Office's research findings to a limited number of educational researchers and other interested persons.

OFFICE OF RESEARCH STAFF

Don A. Abshire, *Programmer*

Richard K. Anderson, *Programmer*

Alexander W. Astin, *Director*

Alan E. Bayer, *Associate Director*

Barbara A. Blandford, *Project Director, Higher Education Panel*

Judy M. Cangialosi, *Secretary*

John A. Creager, *Research Associate*

David E. Drew, *Research Associate*

Jeffrey E. Dutton, *Chief, Data Processing*

Jerome B. Karabel, *Research Assistant*

Melvena W. Kelley, *Secretary*

Laura P. Kent, *Editor*

Margo R. King, *Secretary*

Deborah Kurzberg, *Secretary*

Terry G. Mahn, *Programmer*

Linda D. Molm, *Research Analyst*

Jack E. Rossmann, *Research Associate*

Jeannie T. Royer, *Secretary*

Charles C. Sell, *Associate Chief, Data Processing*

Joan C. Trexler, *Research Assistant*

Richard M. Webb, *Programmer*

ACE RESEARCH ADVISORY COMMITTEE

N. L. Gage, *Chairman*
Professor of Education and Psychology
Stanford University

Jessie Bernard
Research Scholar Honoris Causa
Pennsylvania State University

John G. Darley
Chairman, Department of Psychology
University of Minnesota

James A. Davis
Director
National Opinion Research Center

Edgar G. Epps
Professor of Urban Education
The University of Chicago

Richard C. Gilman
President
Occidental College

Edmond P. Gordon
Professor and Chairman, Department of Guidance
Teachers College
Columbia University

W. Lee Hansen
Chairman, Department of Economics
The University of Wisconsin

William Sewell
Professor, Department of Sociology
University of Wisconsin

Michael Useem
Assistant Professor, Department of Sociology
Harvard University

Dael Wolfle
Graduate School of Public Affairs
University of Washington

Additional copies of this *Research Report* (Vol. 7, No. 2, 1972) may be obtained from the Office of Research, American Council on Education, One Dupont Circle, Washington, D. C. 20036.

MEASUREMENT ERROR IN SOCIAL AND EDUCATIONAL SURVEY RESEARCH

Robert F. Boruch

John A. Creager

Office of Research

American Council on Education

ACE RESEARCH REPORTS

Vol. 7, No. 2

May, 1972

Table of Contents

		Page
	List of Tables	
	Acknowledgements	
I	Introduction	1
II	The Meaning of Measurement Error	2
III	The Effects of Error on Analysis and Interpretation	5
IV	An Overview of Pertinent Research	7
	A. Generation of Error as a Function of Item Type	7
	B. Generation of Error in Data Processing	10
V	The Reliability of Freshman Survey Items	12
	A. Sampling, Administration, and Processing	13
	B. Stability of Response as a Function of Item Content and Form	14
VI	The Stability of Self-Reported Protest Activity and Attitudes	19
VII	The Accuracy and Precision in Self-Reports of Grades and Age	22
	A. Self-Reports of Freshman Grades	22
	B. Self-Reports of Age	24
VIII	Errors in Initial Data Files: Verified Keypunching Vs. Optical Scanning	25
IX	Concluding Remarks	27
X	Bibliography	31
XI	Appendices	
	A. Appendix A: 1969 Student Information Form	51
	B. Appendix B: Supplemental Student Information Form.	57

List of Tables

Page

1. Test-Retest Statistics and Reliabilities for Items on Demographic Characteristics, Family Background, High School Achievements, and College Admissions..... 39
2. Bivariate Distribution for Stability of Reported Father's Education..... 39
3. Bivariate Distribution for Stability of Reported Mother's Education..... 39
4. Bivariate Distribution for Stability of Estimated Parents' Income..... 40
5. Bivariate Distribution for Stability of Reported Concern About Financing College Education..... 40
6. Bivariate Distribution for Stability of Reported Average High School Grade..... 40
7. Bivariate Distribution for Stability of Reported Rank in High School Class..... 41
8. Bivariate Distribution for Stability of Reported Distance of College from Home..... 41
9. Bivariate Distribution for Stability of Reported Region of Birth..... 41
10. Test-Retest Statistics and Reliabilities for Four Reported Sources of Financial Support During Freshman Year..... 42
11. Test-Retest Statistics and Reliabilities for Reported Attributes of Parents..... 42
12. Test-Retest Statistics and Reliabilities for Reported Chances of Future Events..... 43
13. Test-Retest Statistics and Reliabilities for Reported Objectives..... 43
14. Test-Retest Statistics and Reliabilities for Reported Attitudes Toward Federal Involvement in Problem Areas... 44
15. Test-Retest Statistics and Reliabilities for Reported Attitudes about Campus and Social Issues..... 44
16. Test-Retest Response Probabilities and Phi Coefficients for Checklist of High School Achievements.. 45
17. Stability of Reported Current Political Preference..... 45
18. Stability of Expected Political Preference in Four Years. 45

List of Tables (Con't.)

		Page
19.	Stability of Reported Participation in Protest Against Racial, Military, or Administrative Policy During Senior Year in High School.....	46
20.	Stability of Freshmen Predicting Future Involvement in Protests about Racial, Military, and Administrative Policy.....	46
21.	Stability of Reported Attitudes on Items Pertaining to Student Freedom and Administrative Control.....	47
22.	Agreement Between Self-Reported and Recorded Freshman Grades (Proportions).....	48
23.	Agreement Between Self-Reported and Recorded Freshman Grades of Black Students.....	48
24.	$\emptyset$ and $\emptyset/\emptyset_{max}$ for Self-Reported and Recorded Freshman Grades, by Sex and Race.....	49
25.	Self-Reported and Computed Proportions of Students in Age Categories.....	49
26.	$\emptyset$ and $\emptyset/\emptyset_{max}$ for Self-Reported and Computed Age.....	49
27.	Correlations Between Item Responses Processed by Optical Scanning and by Verified Keypunching.....	50

Acknowledgements

This study was supported by NIMH Grant 1 R12-MH17 and NSF Grants GR-89 and GS320-73X. The professional staff of the American Council on Education's Office of Research made many helpful suggestions in its planning and execution and in the preparation of this report. Margo King and Mel Kelley provided secretarial assistance. Jeannie Royer prepared the final manuscript and supervised arrangements for its printing. Laura Kent made many useful editorial suggestions on the draft manuscript.

We are especially grateful to administrative personnel at the participating institutions for their help in arranging time and space, for administering the survey forms under the proctorship of the authors, and, of course, we appreciate the voluntary participation of the students who provided us the data for the study.

Measurement Error in Social and Educational Survey Research

Robert F. Boruch[1] and John A. Creager

With the development of large-scale longitudinal research programs in the behavioral sciences, the use of omnibus questionnaires to solicit data has greatly increased. Typically, such questionnaires reflect extensive and legitimate concerns about the impact of social, cultural, and educational processes on the individual. To determine causal and other functional relationships, data relevant to inputs, to treatments or environments, and to outcomes are collected and then subjected to analytical procedures, often of a multivariate, correlational nature. In addition to their descriptive value, these data are being used more frequently as a basis for appraising policy and making management decisions.

Data from survey questionnaires are, however, subject to measurement error, potentially a major source of inferential errors and, therefore, of erroneous recommendations based on the data. Although the nature and effects of measurement error have been extensively treated in some areas of social and behavioral research, and empirical estimates of reliability of data are available for a variety of psychological measurements, there is little empirical information about measurement error in the data obtained by survey questionnaires which can be feasibly administered to a heterogeneous sample of widely dispersed subjects. Even when available, such data are seldom explicitly related to the relevant theory of measurement error, its meaning and effect on analytical inference. This situation is hardly tolerable in view of the use made of such information in reaching important policy decisions.

The Cooperative Institutional Research Program (CIRP) of the American Council on Education is an example of a systematic program of social and educational research in which survey methods using questionnaires constitute a major technique for collecting data. This longitudinal research program, described in greater detail elsewhere (The ACE Office of Research: Its Purposes and Activities, 1972), is dedicated to the study of higher education, particularly its impact on student development. In such a program, measurement error may affect data analysis and interpretation at each stage of the research system through which data are transmitted and

---

[1]Now with the Department of Psychology, Northwestern University.

processed. That is, inaccuracy or imprecision in recorded response, and in the statistics based on these data, may occur not only during the solicitation of data, but also during information transcription, and file development. Moreover, many factors determine the nature and extent of errors: the substantive nature of the solicited information, the physical and administrative conditions under which the data are collected, and the characteristics of the relevant groups involved: e.g., the respondents, the researchers, and the users of the information.

The purpose of this report is to present empirical results about certain problems of measurement error in an educational survey system. These results are confined to measurements derived from items in the annual surveys of entering freshmen that are a basic part of the CIRP. The survey information serves as input data for the longitudinal studies, and as the basis for published national norms on freshmen entering American institutions of higher education (Creager, Astin, Boruch, Bayer, and Drew, 1969).

After discussing the meaning and consequences of measurement error, we present a review of the pertinent literature, and then give the results of recent empirical investigations which we have conducted. The discussion is restricted primarily to one method of measurement, the survey questionnaire designed for efficient processing in large volume, and to one type of index of reliability, the temporal stability of item response.

The Meaning of Measurement Error

Measurement error is defined as the difference between a recorded response to an inquiry and a potentially measureable, true condition associated with that inquiry. Some of the sources of measurement error in a survey questionnaire item are the respondent's faulty recall, his deliberate or accidental distortion of response, structural weakness or semantic ambiguity in the item, and lapses in the quality of data recording, transmission, and maintenance. Measurement error can be defined more precisely in terms of actual measurement operations.

One starts with some conceptually definable attribute or condition of an observation unit (e.g., an individual subject) and some specified operations purported to measure that attribute or condition. Empirical evidence shows that the resulting observed score is subject to variability when the operation is repeated. On the assumption (usually tacit) that the attribute or condition has not, in fact, changed, an observed score consists

of a variable error component in addition to the constant value or "true score" which an unchanged attribute should have yielded from the measurement operation. Unfortunately, no operation exists for partitioning a single observed score into independently determined values for the true score and the error components. In rare circumstances, the attribute or condition can be determined independently, objectively, and accurately; the subject's response may then be compared with the actual condition, in which case we can speak meaningfully of veracity of response.

In the absence of operations for directly partitioning of a single observed score into true and error components, operations have been developed for partitioning the variances of these components, these variances being computed over defined groups of observation units. To accomplish this, we must introduce the assumption that, when averaged over a group of observations, individual errors tend to cancel each other out yielding a zero mean, and that any statistical dependencies between errors and true scores also cancel out, yielding null covariances between the true score and error components. Because much of the methodology for data analysis involves variances and covariances, this formulation has proved eminently practical for estimating measurement error and for studying its effects on other statistics.

In this formulation of the problem, that an observed score consists of true score plus error, the assumptions discussed above permit the error term to be regarded as a random variate, even though at the level of an individual measurement, the sources of measurement error may be, and usually are, nonrandom in their occurrence and, therefore, in their effects on the observed score. Unlike the situation in the classical physical science where the chance speck of dust got into the control mechanism of a telescope, or into a crucible during chemical analysis -- situations readily minimized by careful control of experimental conditions -- some of the sources of measurement error in surveys, even where known, cannot readily be controlled by the investigator. To be sure, he can, and should, construct his instruments and use survey logistics which either reduce the likelihood of such errors or randomize the errors across observation units. Failing this, even the most elaborate theory and operations for coping with the effects of measurement error can only partially protect against inferential errors.

Given the basic formulation of measurement error as "variability on replication," a variety of experimental operations for further defining and

operationalizing concepts of measurement error can be derived. At issue is the nature of the replication. In classic psychometry, replication means subjecting the same observation units to a set of items presumed to measure the same attribute or condition; the interitem (or interjudge) agreement as measured by item covariances, is then used to assess the internal consistency of the item set. If a composite score is derived by summing responses to items in the set, and if the item-composite covariances are high, the set may be regarded as scalable. (Note that the pattern of item difficulties is also involved and affects the variance-covariance pattern in the set.)

In many questionnaires, one is dealing not with a scalable set but with single items in a checklist or with specific questions (e.g., "What is your sex?"). In this situation, the internal consistency of an item set cannot be used to estimate measurement error. We need another kind of replication. When the same measurement operation is repeated at a later time, under conditions as similar to the first as possible, agreement on replication is called temporal stability. If observed change in a score overtime is to be regarded as measurement error, it is necessary to construct the operations such that the true attribute or condition has not in fact changed. The temporal stability concept applies both to responses to single items and to scores derived from scaled item sets. It can thus be applied to survey questionnaire data. Unfortunately, no operations exist for ascertaining that optimal retest time-lag which sharply separates changes in true score from measurement error.

Finally, replication may consist of using two or more ostensibly different measurement operations, both of which purport to measure the same attribute or condition in question. The operations may differ either in the measuring instruments used (alternate forms) or in the experimental conditions of their administration. Because this third kind of replication, called equivalence, necessarily involves a time lapse, some confounding with temporal stability is inevitable.

These various approaches to ascertaining measurement error based on replication of measurement of the same attribute or condition, are subsumed under the generic term reliability. When two measurements of presumably different attributes or conditions are compared, we normally speak of validity. So-called face validity results from comparing a measurement operation with an a priori concept of an attribute. With this exception, both

reliability and validity are ascertained by correlations. If correlations are high, indicating measurement of the same attribute or condition, we usually think in reliability terms with lower, or moderate correlations being thought of as "validities." Any degree of uniqueness suggests that two measurement operations are not measuring the same thing, while any degree of communality less than unity suggests imperfect measurement of the same or similar functions. In practice coefficients may range from -1 to 1 so that a sharp empirical distinction between reliability and validity is not as convincing as the purposeful conceptions of scientists would lead us to expect. Nevertheless, the classical distinction between reliability and validity has proven to be a useful one.

The Effects of Error on Analysis and Interpretation

Whether one is estimating descriptive parameters or testing hypotheses about treatment effects, the major reason for concern about measurement errors is that they may affect any inferences drawn from the data. These inferences occur in analysis and interpretation and may ultimately damage the credibility and relevance of the reported results and recommendations to users of the information. The effects of measurement error -- and, therefore, appropriate ways of coping with them -- differ somewhat according to whether one is dealing with continuous or multinomial data. We will first consider error in continuous data.

A recent AERA Symposium (1971)[2] emphasized some of the risks in correlational analysis and dealt with ways of coping with these risks in the case of continuous variables (Finucci; Gleser; Stanley and Livingston; Edwards; Porter). These papers review, criticize, and elaborate on classical attenuation theory, thus providing a useful approach to the problem. Astin (1970) makes explicit some effects of measurement error on substantive inferences drawn from analyses. Writing about the methodology of college impact research, he expresses special concern about the effects of measurement error in input variables on the input-output regression, pointing out that such errors may result in overestimating college impact when examining the output residualized on the input variables.

Different patterns of measurement error across input (pretest) and treatment variables may result in an effect opposite to and confounded with the

---

[2]The symposium, "Some Attenuating Effects of Errors of Measurement," was organized by Julian C. Stanley and chaired by Gene V. Glass.

one discussed by Astin, in which case we may fail to detect controlled output differences and impacts when they actually occur. This point may be illustrated by the hypothetical case where a student input variable with reliability of .95 based on mean observations within a college is correlated .5 with a college characteristic variable with reliability of .60. If the true regression weights for predicting some outcome were both equal to 2.0, the regression weights estimated from the fallible data would be 1.47 and 1.05, respectively. Not only is the importance of <u>each</u> variate underestimated, but also the student effect appears to be more important than the college effect. Moreover, the order of magnitude of the regression on weights may be completely reversed. Using formulas given by Cochran (1968), for example: if "true" regression coefficients were 2.0 for college effects and 1.0 for student effects, with reliabilities of .60 and 1.00, respectively, the estimated regression weights would be 1.15 and 1.25. This is not to say that previously reported results indicating student input to be more important than college effects are necessarily wrong, but rather that inferential errors may occur in either direction and that in a given analysis, measurement errors may have multiple effects which render the interpretation uncertain.

Turning to misclassification error in multinomial data and contingency tables, Assakul and Proctor (1965) and Cochran (1968) provide scholarly treatment of this subject. With such data, the dependence of variances on means -- and therefore of errors in estimating variances on errors in estimating means -- makes the issue of how measurement error affects inference more complex than in the case of continuous variables, but we can say that, when data are fallible, true differences between groups or between the same group measured at different points in time will be more difficult to detect. One may appraise the credibility of an obtained percentage or Phi coefficient by estimating the misclassification rate and computing a statistic "corrected" for the error; a computer program for obtaining tabulated values of maximum Phi for a variety of misclassification rates is given by Boruch (1972, in press).

These considerations emphasize the need for obtaining reliability estimates not only for student input variables but for environmental and outcome (dependent) variables as well. It is beyond the scope of the present study to provide such estimates, though a few comments are in order. The reliabilities of many college environmental variables have been reported by researchers,

like Astin, Holland, and Pace, who are concerned with assessing the college environment. Moreover, many of the outcome variables are retests of input variables, and therefore it is plausible to assume their reliabilities are similar to those on pretest. Nevertheless, data are needed to check this assumption and to provide estimates for those outcome variables which cannot be pretested.

An Overview of Pertinent Research

A number of writers in the scientific and educational research literature have considered measurement error -- e.g., Kish (1967), Cochran (1968), Parten (1950), and Keating (1950) -- but few have presented quantitative data on the topic. A recent exception is the presentation of data on self-reported high school grades and noracademic achievements (Maxey and Ormsby, 1971). Nevertheless, Bradburn (1969) and Siegel and Hodge (1968), among others, have outlined the weaknesses of current survey research, citing both the paucity of data and the inappropriateness of classical statistical models in dealing with fallible data.

Although many journal articles report research based on fallible data, little or no attention is given to appraising empirically the extent of error or its impact on inferences. Menne, Klingensmith, and Nord (1969), after systematically examining articles in all issues of six widely circulated educational research journals published over a recent five-year period, concluded that, in at least half of the articles, insufficient attention was paid to the accuracy and precision of the cited measures. In only one journal did the number of articles that provided estimates of reliability of measurement exceed the number in which such estimates should have been made but were not.

In the case of survey data the lack of reliability data is understandable even though undesirable. The logistic and cost constraints of obtaining empirical estimates of the reliability of survey data can be severe because such estimates require replicated measurements under similar conditions with heterogeneous and dispersed respondents. Nevertheless, existing data are helpful insofar as rough estimates of reliability for various survey item types can make analytical results more credible.

Generation of Error as a Function of Item Type

Astin (1965), Smith (1932), and Walsh (1967, 1968) have developed reliability estimates for a series of biographical items, including questions with factual

referents. Smith's early investigations of the stability of college students' responses to questionnaire surveys used a sample of only 50 students and a period of four months elapsed between the initial and second surveys. Questions about the number of courses the student had failed resulted in the most stable responses (91-97 percent concordance) of all the questions pertaining to academic performance, whereas estimates of high school size and reports of high school extracurricular activities changes the most from test to retest (50-80 percent concordance).

Smith also reports interesting sex differences in the stability of reports of family attributes. Estimated parental income was changed on retest by 36 percent of the women and by 46 percent of the men. Women were more variable in their answers about their fathers' occupations (11 percent change) than were men (5 percent change), but 13.6 percent of the men changed their response to a question about their mothers' occupations in contrast to only 4 percent of the women being inconsistent.

Low stability was characteristic of responses to personal value items: least preferred extracurricular activities, favorite books, and favorite authors or historical figures. The highest stabilities were associated with queries about liking people, liking the opposite sex, and getting along with siblings; for all these items, the change in response was less than 20 percent. Although the small sample size makes these comparisons extremely tenuous, there is some confirmation of the results in other studies.

Astin (1965) administered a questionnaire twice, with six weeks intervening, to a sample of 107 college students; questions on previous achievements resulted in 90-100 percent agreement (binomial classes of response) between survey and resurvey responses. Responses to items characterized by little ambiguity and by the "importance" of the accomplishment (e.g., "elected student president") had high rates of stability (95-100 percent). Achievements of a more ambiguous sort (e.g., "received an award for leadership") resulted in slightly smaller rates of agreement (90-95 percent).

Responses to inquiries about other facts -- e.g., father's education and occupation, the student's average high school grades, and certain nonacademic accomplishments -- were subject to greater variability: The percentage agreement from survey to resurvey has a range of 74-92 percent. Attitudinal items and questions which solicited information on future plans (e.g., probable major field of study) typically involved concordances in the 60-70 percent range.

Walsh (1967, 1968) has investigated the relative accuracy of data from questionnaires, interviews, and biographical inventories in soliciting verifiable information from college students; he also provides a thorough review of pertinent research reported in the psychological literature. His inquiry was limited to questions pertaining to academic grades, major discipline, and number of semester hours completed. Forty-five students randomly selected from men's residence halls in a large midwestern university were allocated to each of the three methods of data solicitation; their self-reports on grades, major fields, and semester hours completed were checked against administrative records. The results suggest that the method of soliciting information does not substantially affect the veracity of response to these limited types of questions. The self-reports about the number of courses dropped and the number of courses in which the student received a "D" or failing grade were most accurate (80-100 percent). Least accurate were self-reports on cumulative or recent semester grade-point averages (50-80 percent) and retrospective reports of high school grades and rank in high school class (51-53 percent). Again, because the sample size was small, these results are suggestive rather than definitive.

Unfortunately, little is known about differential variability as a function of sex, race, and other demographic characteristics of respondents, although some research (e.g., Bailar, 1968; Hansen, Hurwitz, and Pritzker, 1964) suggests that such differences in response stability do exist. It is also plausible to expect that covariances in the response variations exist. Functional interdependence is certainly important insofar as many common techniques for accommodating measurement error assume that the variations are independent of one another and of the responses to questions about attributes under examination (Cochran, 1968).

Further documentation of the statistics on measurement error is available insofar as many factual items in educational survey questionnaires overlap considerably with information solicited by the U.S. Census Bureau. These questions include age, sex, parental education and occupation, religious preferences, and income levels. A prime basis for evaluating the survey-resurvey reliability of such items is furnished by Hansen, Hurwitz, and Pritzker (1964). The work of these researchers and of Bailar (1968) and Haber (1966) typify many of the investigations in this area, where simple variances of census interview-reinterview responses and other error parameters are estimated for

subgroups of the national population. Although only part of the information from these studies is relevant to the college population, the statistics on variability reveal differential variability across subgroups for a specific attribute: Lower-income families are more likely to be inconsistent in their reports of income from one survey to the next than are upper-income families, even when the resurvey follows with a short time lag.

Generation of Error in Data Processing

The transcription of individual responses to a machine-readable mode (i.e., card or tape) is a second stage at which error may be introduced into the data system. Here, two broad areas are germane to large-scale surveys: the veracity of clerical transcription and the veracity of electromechanical transcription.

Minton (1969) has perhaps done the most extensive work in the domain of clerical processing. Although his research is based on the U.S. Census procedures, the results are pertinent to other surveys. Minton discusses the detection and correction of defective data transcription from document to card mode, emphasizing quality control paradigms for classifying the nature, source, and extent of such error. He also discusses key-punching, with dependent and independent verification, clerical coding, and proofreading of typed documents. Relying primarily on research conducted in 1950 and 1963 and involving small samples (1000-3500) of cards purposely inoculated with errors, Minton found that 3-4 percent of the defective items were undetected in 100 percent verification of punched cards. Reverification of already verified cards yielded somewhat different error rates: 14 percent for a sample of 5,541 cards; 12 percent for a sample of 3,314 cards. The discrepancy between the two sets of error rates is attributable, in part, to verifiers learning to recognize errors deliberately implanted in cards. Since the coding of responses to questionnaires is more influenced by interpretation and judgment, it can be expected to involve even higher rates of error and lower rates of error detection. Minton found that, with this type of operation, 30-70 percent of defectives were corrected in sample verification.

The accuracy of commercial mark sensing devices (mechanical, electrical, or optical) when used in survey questionnaire processing has not been treated extensively in published research, although occasional technical memoranda and anecdotal evidence are available. Such reports are generally characterized

by interest in the number of machine readable marks made on a document, as compared with the total number of marks on a document. Studies by Valentine (1959, 1960) concern the effect of light, medium, and heavy pressure on the sensing capability of a mark sensing device (a machine currently used in commercial and educational operations). A single mark in response to a query is made by the respondent on each sector of a mark sense card. A defective card was defined as one where the marks were too light for sensing in some sectors or where a single sector contained multiple marks. The rejection of defective cards was approximately 33 percent for light marking, 11 percent for normal pressure, and 3 percent for heavy pressure (sample size of 428 with five sectors per card). The use of mechanical rather than wooden pencils resulted in near-zero rejection rates.

Frequently, commercial and research organizations that require large-scale data processing conduct accuracy tests on mark sensing devices prior to purchasing such equipment. The results of such tests are rarely published, however, so relevant data are sparse. One exception is a study by Spencer (1966) of the reliability of a highly standardized DIGITEK test-scoring device. The results provided by the machine agreed highly with error-free results derived by hand-scoring with verification: Reliability in scoring was 99 percent with 99.4 percent stability (i.e., overlap in results of repeated scoring of the same protocols by machine).

Personal communications between clients and vendors of different equipment are more typically the mode of information exchange. For the survey questionnaire considered in this report, the reliability of scoring by optical scanning is said to be in the range 92-99.5 percent. The most recent data on the optical scanning equipment used suggests that reliability is near 100 percent when the ratio of dark or heavy pressure marks to light marks and omissions is high. As the proportions of light marks on an answer form increases, the sensing capability becomes less accurate (Dillon, 1970).

Computerized editing can be used to detect and correct defective data generated at earlier levels of the research system. This procedure is typically used to discover errors and to correct the data for further analysis. Processing is directed toward detecting inconsistencies and contradictions in the data in cases where external and internal criteria can be employed to define consistency. If one knows the total frequencies and percentages, the output from data processing can be monitored; the range of allowable responses

can be checked to insure that no gross errors are committed; the extent of missing data may also be examined. Internal consistency is typically determined by checking certain internal counts against alternative criteria. Moreover, sample data with known errors (i.e., a "hot deck") may be used to identify and correct internal operations which function as devices for quality control. Freund and Hartley (1967) present examples of such procedures and describe flow charts for computer programs that may be helpful in editing the data. Their methods of inferring erroneous data depend on the logical consistency of the relations among data, and on the empirical reliability and variability of the defect-free information in the data file. Examples of similar algorithms and programs for computer editing are furnished occasionally in Datamation, Educational and Psychological Measurement, and other technical journals on computing machinery and data processing. O'Reagan's (1969) analysis gives evidence for the cost benefits of similar automatic data editing procedures when the alternative procedure is manual correction.

The Reliability of Freshman Survey Items

As part of the Cooperative Institutional Research Program (CIRP), each annual cohort of freshmen entering participating institutions completes a four-page survey questionnaire. This instrument is designed to elicit responses to a broad range of items presented in multiple choice of checklist formats, and requires about 25 minutes to complete. The completed questionnaires are readily processible by an optical scanner, which senses the respondents' answer marks and records them in coded form on a magnetic tape for further analysis with an electronic computer. Such procedures enable the processing of data from a nationwide sample of approximately 300,000 respondents each year. About one-third of the items are repeated from one survey to the next, another third appear less often, and the final third are new in a given survey. This arrangement provides both continuity for the assessment of trends and flexibility for the exploration of current issues in higher education.

The pool of items used in a given year solicits information on the demographic and biographical characteristics of entering freshmen, their high school experiences and activities, their academic and career plans, their life goals and aspirations, and their attitudes toward various political and academic issues. Since our aim in this study was to obtain estimates of the reliabilities of all freshman survey items used in questionnaires for the four

cohorts -- 1966, 1967, 1968, and 1969 -- the 1969 Student Information Form (SIF) was supplemented with a special form consisting of those items in prior surveys which had not been repeated in the 1969 form. A copy of the 1969 SIF appears as Appendix A; a copy of the supplemental form appears as Appendix B.

Sampling, Administration, and Processing

Although we had originally planned to conduct the study with a subsample of institutions participating in a regular survey, and to use the actual data from this survey as the pretest data, thus making the retest as similar as possible to actual survey conditions, this plan proved to be impractical financially and logistically. Therefore, we administered the two questionnaires (the 1969 SIF and the Supplemental SIF) to second-term freshmen at three institutions, none of which were participants in the 1969 survey: the groups were retested two to three weeks later. The three institutions included a state university (N=97), a private university (N=62), and a community college (N=43) located in the metropolitan area of Washington, D.C. Matching test-retest protocols were obtained from 202 students; however, response rates for individual items ranged from 195-202. All but three of the students were Caucasian, thus precluding any systematic study of racial differences in stability of response. Students who completed both the test and the retest questionnaires were paid a small honorarium for their participation. Subjects were obtained through institutional cooperation for making testing facilities available and by informing students of the study through posted notices and announcements in the student newspaper.

The sample was small and somewhat biased, and the testing was carried out after the freshmen had been in college for more than one term, in contrast to our usual practice of surveying entering freshmen during the registration or orientation period. Nevertheless, these conditions were probably adequate to ensure useful estimates of item reliability. The sample can be characterized more fully by noting that the institutions participating in this study were neither very selective nor very affluent. The subjects were predominantly white students from the urban areas in the eastern and southern regions. Compared with the national population of freshmen entering college in 1969 (Creager, et al.), the students participating in this study had typical high school grades, but a larger proportion ranked in the top quarter of their high school graduating classes; in addition, their parents were somewhat more highly educated and the family somewhat more affluent.

Stability of Response as a Function of Item Content and Form

Estimates of the reliability of survey items obtained from the 202 freshmen in this study are presented in a series of tables, of which there are two major types: in the first, distribution statistics for both test and retest, as well as the reliability coefficients are presented for blocks of items similar in form or content; in the second, the bivariate (i.e., test-retest) percentage distributions are presented for selected single items. The first type enables easy comparison of reliabilities for items covering a given area or presented in a certain form. The second type illustrates the actual deviations from retest regression and variations across item categories; rows are defined by pretest categories and columns are defined by posttest categories. Generally, these bivariate tables show that the test-retest regressions are rectilinear and homocedastic.

The test-retest statistics and reliability coefficients for demographic characteristics, family background, high school achievements, and college admissions items (presented in Table 1) proved to have the expected high stability. Most of the reliabilities were so close to 1.0 that corrections for measurement errors are probably not needed when such variables are used in analysis. One exception was the item dealing with the student's concern about financing his college education, which had an r of .85. This item and the item on age are the only ones in the block for which the true score can change with the passage of time; any response variability in the other items is probably attributable to faulty recall or to carelessness in reading and marking the questionnaire. When the respondent is asked to indicate his age as of a specified date, as he was, only the means and standard deviations should change at that date. When he is asked to give his age as of the date of testing, or at his nearest birthday, more short term variability in response is to be expected. Concern about financing one's education can, however, alter at any time. This item has three alternative response categories; the other items in this group have from six to nine categories. Similar reliability estimates (.85-.88) were observed for the block of items on sources of financial support during the freshman year (Table 10), which also have three alternatives.

The bivariate plots shown for these items in Tables 2-8 are just what one would expect given such high reliabilities; the few deviations observed between test and retest data were largely confined to shifts to an adjacent category. For such variables as reported annual income of the parents,

appreciable skewness in the distributions, similar in degree and direction, appeared both on test and retest: This is apparently a characteristic of the true score distributions rather than the result of random sampling errors. The bivariate plot for the region of birth, generated from the student's reported state of birth, is shown in Table 9. The overall test-retest concordance for the region categories is 98 percent, the values for individual regions ranging from 96 percent to 100 percent.

Reliability data are not presented for either the sex or the racial/ethnic items. No valid estimate of the reliability of the sex item could be obtained for the total sample because of an experimental error.[3] There is, however, no reason to believe that the sex item would be any less reliable under normal conditions of questionnaire administration than were the other demographic variables. Any response variation on this item is most likely to be a function of sporadic carelessness in marking the survey form.

In the case of the racial/ethnic item, no estimate of reliability was possible because only three blacks were included in the sample, and other racial/ethnic groups were not represented at all. Investigating the reliability of this kind of item and particularly its sensitivity to the number and semantic form of the response alternatives, Bayer (1971) found that a properly constructed item can have a short-term retest concordance of 98 percent or higher. If carelessly constructed, however, such an item may result in very serious errors in identifying racial/ethnic subgroups; any special analyses or comparisons based on such fallible data can lead to very misleading inferences. The item is not one on which the true score can vary, if the racial/ethnic categories have any absolute meaning beyond the culturally determined semantics of self-identification. Hence, response variability is primarily a function of item structure, over which the researcher has some degree of control and, therefore, responsibility.

In the area of academic aspirations (e.g., degree plans, probable major field, and career plans), short-term retest resulted in nearly perfect concordance. The sporadic examples of discordant response appear to be attributable

---

[3] At two of the institutions, students were told that they could use aliases, provided that they were the same on both the test and retest forms. Of those who did so, some used names associated with the opposite sex and were inconsistent in whether they checked their true or fictitious sex. The estimate of reliability for 60 students in the third institution, where real names were used was .96.

to careless marking of the survey form. Considerable changes have been reported in the response to these items over the longer period of the freshman year (Bayer, Drew, Astin, Boruch, Creager, 1970) and these changes differed according to the type of college attended. In view of the high concordance of response in the present reliability study, the changes observed over the entire freshman year are very convincing and are most likely to be a function of the college experience.[4]

With the exception of the concern for financing one's education, and reported sources of financial support, the reliabilities of the items considered so far were sufficiently high that unity may be reasonably placed in their associated diagonals in correlation matrices. Even though no variable is perfectly reliable, it should be recalled that corrections for attenuation involve the square roots of the reliabilities and that the values obtained for the reliabilities are themselves estimates subject to sampling errors. For the remaining variables considered in this study, reliability estimates should be used in the diagonals of correlation matrices for regression or components analysis.[5] In some cases, such replacement of diagonal values may result in some side correlations of unity or greater, in which case one of the offending variables should be eliminated, or the highest uncorrected side correlation used as a better estimate of reliability. In this connection, computer software permitting a test of matrix singularity, rearrangement, or elimination of variables may prove useful (Boruch and Dutton, 1970).

---

[4]In the case of dichotomous variables generated from the responses to a single category, the reliability as measured by the Phi coefficient is a function of the base rate, i.e., the frequency with which that category is checked. This, in turn, is not independent of the number of, and base rates for, the other categories. Hence, the Phi coefficients may be low even when there is a high concordance on retest. The Phi/Phi max coefficients, which allow for the variation in base rates, are much higher and more consistent across categories, but have the disadvantage of representing an intrinsic relationship rather than one which is practical in correcting for attenuation in correlational analysis (Guilford, 1950).

[5]Communality estimates should be used for common factor analysis. The equivalent replacement of unit diagonal values may be accomplished by pre- and post-multiplication of the observed correlation matrix by a diagonal matrix containing the replacement values (reliabilities or communalities), prior to the correlational analysis. Some generalized regression programs for computerized analysis include this option. Many, however, do not and thus do not function properly without unit diagonals.

In contrast to the high reliabilities so far noted, responses to items on attitudes and opinions, life goals, and the estimated chances of future behaviors were considerably less stable (Tables 11-15). When students were asked to report on certain attributes of their parents, the reliabilities were in the range of .60-.82 (Table 11). One likely source of this response instability -- beyond the fact that the student was asked about "what is inside the head (or heart)"of others -- is that the attributes as perceived may be different with respect to the father and mother; in this case, what does the respondent do, when asked to characterize his parents as a unit: report for one parent on test and the other parent on retest? Or give a differently weighted average on each of the two occasions? Asking for a single response about both parents is faulty item construction; either ask about one specified parent, or allow separate responses for each parent.

Reliabilities for students' estimates of the probability of certain future events occurring, (Table 12) ranged from .58 to .88. The response stability was clearly a function of the event: e.g., unstable for expecting to drop out of college temporarily, but fairly stable for expecting to transfer to another college, possibly because two-year students who plan to go on to a four-year institution were included in the sample. Response stability was lower for those events which are dependent chiefly on the actions of others: e.g., being elected to a student office, getting an article published, graduating with honors. It was highest for getting married while in college, joining a fraternity or sorority, and for protesting over various issues, all actions that are more under the control of the respondent. The reliability of the protest items is considered in greater detail in a later section of this report. The reliabilities of students' reports of the life objectives they considered to be essential or very important (Table 13) ranged from .65 to .87 and again varied with the nature of the specific objective. They tended to be higher for such goals as raising a family, being well-off financially, keeping up with political affairs, and being creative in the arts. The lower value for developing one's life philosophy is probably a function of its high popularity (i.e., base rate). In this connection it should be noted that items which have lower stability as a function of extreme base rates may nevertheless be important and useful in research.

Attitude and opinion items have lower average response stabilities and more variable estimates across items, both for attitudes toward Federal

involvement in various social and economic areas (Table 14) and their attitudes toward various academic and social issues (Table 15). Many freshmen enter college with limited awareness of, and information about, or even interest in, many of the issues. Some, perhaps, remain uninterested and uninformed; many other students may develop rapidly in these respects, especially during their early encounters with other students from widely different backgrounds. Thus, despite their modest reliabilities, attitude items have considerable value for a longitudinal research program in which college experiences are hypothesized to help a person become more aware of, and better informed about, such issues. Given this degree of instability responses of freshmen to these items, it is necessary to make an independent estimate of the reliability of their responses to these items when they are reported in the followup surveys. Only if this is done can the college effects on reliability be separated from effects on mean attitude changes.[6]

Phi coefficients for checklist items of secondary school achievements (Table 16), scored as dichotomies, were uniformly high, despite low base rates for most of the items on the list. For two of them, participating in a National Science Foundation summer program and winning in a state or regional science contest, the base rate in the sample was too low even to permit an estimate. The relatively high reliability of the remaining items probably resulted from asking the student for objective information about his past in contrast to asking him for subjective judgments or about future plans and expectations.

Freshman reports about their current and future political preferences as indicated on a scale of five ordered categories were extremely high and nearly identical in their concordance patterns (Tables 17-18). Here we do not observe the drop in reliability usually associated with future-oriented items. In this case, other available data show that there is an actual shift toward the political "left" during college, both within and between class years.

---

[6]The reliability of responses to followup surveys conducted by mail is important in the analysis of panel data and generally more difficult to estimate than in the case of the freshman surveys. For many of the items repeated in followup surveys, the estimates from the freshman survey would seem to be reasonable ones to use in the absence of independent estimates. This notion becomes more tenuous, however, the lower the estimate from the freshman survey, and, as noted above, can be misleading in the case of items on attitudes, opinions, and expectations.

(For several supporting references, see Bayer, Astin, and Boruch, 1971.) Again, the high reliability of the input data strengthens the conviction that these longitudinal changes are changes in true scores, not artifacts of measurement error.

The Stability of Self-Reported Protest Activity and Attitudes[7]

Because of recent special interest in campus unrest, the Council's Office of Research has not only conducted a number of studies in this area but also has included in the Student Information Form items on protest activity in high school, and on attitudes toward college administrative controls over student life; a self-rating of liberalism has also been included. This section documents the response stability of such items in greater detail, to provide a rough picture of the nature and extent of measurement error in research on protest activities.

Students were asked to indicate whether they had participated "frequently" (coded "3"), "occasionally" (coded "2"), or "not at all" (coded "3") in high school protests against three policies: racial or ethnic policies, military policies, or high school administrative policies. Mean scores and standard deviations for test and retest, along with correlations between test and retest, are presented in Table 19. The columns headed $\Delta\bar{X}_{2-1}$ and $\Delta S.D._{2-1}$ are the algebraic differences between pretest and posttest means and standard deviations, respectively.

Judging from the correlation coefficients, and bearing in mind that the subsamples were small, the freshmen from the public institutions (1 and 3) were most likely to change their reports about participation in a high school demonstration against the school administration; the test-retest reliability was rather low for both the state community college (.29) and the state university (.36). The highest reliability (.59) occurred for students at the private university. Either these students were more certain about the extent of their involvement, or they were simply more consistent in their interpretation of the word administrative.

The most stable responses were elicited by the question about protest against military policies. Reliabilities are quite high, ranging from .76 to .93. The reason for the high reliabilities may be the unique and unambiguous

---

[7]The data and discussion portions of this section have been combined with sampling information from an earlier section and are being published separately (Boruch and Creager, 1972).

character of a protest against the Vietnam war, the draft, and related issues. Protests over racial issues were characterized by an intermediate range of reliabilities, from .54 to .74.

Typically, changes in mean response from test to retest are nonsignificant for the students given the option of using aliases (the state and private university groups). In the community college sample, however, the frequency of reported participation in high school protest decreased systematically, rather than being random (and, thus perhaps, attributable to carelessness in filling out the questionnaire).

Since it seems likely that students will tend to give an innocuous answer when the report concerns behavior that might be used against them, this decrease in reported participation probably indicates that they had second thoughts between test and retest about the risk involved in indicating to outsiders that they had engaged in protest.

Freshmen were also asked to estimate the likelihood of their participating in a protest against each of the same policies at some time while in college. Four alternatives were given, ranging from "very good chance" (coded "4") to "no chance" (coded "1"). The findings are given in Table 20.

The test-retest correlations were systematically higher for all three institutional samples and for all three protest issues than was the case for reporting participation in protest activity in high school. Given the widely different likelihoods of participation (as indicated by the within-college means), this stability is encouraging. Homogeneity was also evident in the test-retest means and standard deviations. Despite some college differences, the difference in test-retest averages for any given type of protest was insignificant. The systematic, if nonsignificant, decrements in average means, suggest that some students in each group tended to change their initial response to a more innocuous one (i.e., less likely to participate in a protest).

Table 21 contains data on a variety of issues that are associated with student freedom and with campus protests. The SIF items on which these data are based included a four-point self-rating of liberalism, a five-point item on Federal control of student activism, and four-point agree-disagree items on statements that should regulate activities of off-campus students, should regulate publications and should maintain veto power over campus speakers; the final agree-disagree item states that college administrators are too lax in dealing with campus protesters.

For a given attitude item, response stability did not differ significantly among the institutions, except for the statement concerning administrative control of activism. The students at the public university were less consistent on this item than were the students at the other two institutions.[8] Assuming the equality of correlations across colleges, the weighted averages of correlations for each item except administrative control of activism are:

Liberalism	$r = .65$	Regulate publication	$r = .57$
Ban speakers	$r = .65$	Regulate off-campus behavior	$r = .52$
Administrative laxity	$r = .65$		

Testing for equality of all these estimates except the liberalism rating, we find that differences among estimates are only of marginal significance. That is, choosing a confidence level of .05 leads to rejection of the null hypothesis but using .025 provides no evidence for rejection. The stability of self-ratings of liberalism was significantly higher than that of the other items. The attitudes toward banning controversial speakers from campus and administrative laxity in dealing with protest appeared to be slightly more stable than did attitudes about regulating student publications and regulating the off-campus behavior of students.

The freshmen in this sample were generally most stable in self-ratings of their liberalism. With one exception, previous participation in high school protests were likely to be reported more erratically (from test to retest) when the issue was administrative policy. Reports about participation in military protests were the most likely to be stable.

Agreement with statements endorsing the college's right to ban extremist speakers, to regulate students' off-campus behavior, and to control student publications, and about lax treatment of student protests was fairly stable, but generally lower than liberalism ratings. With respect to Federal control of student activists, the degree of response stability seems to depend somewhat on the particular institution attended by the student.

These data may help to explain certain anomalies in the recent research literature. For example, in a study using identical questions and similar

---

[8] Snedecor and Cochran (1967) provide a test of the hypothesis of the equality of sample correlations. For the statement concerning administrative control of activism, $X^2 = 11.7$ is significant at the .025 level.

groups of students (N = 4,000), Astin (1968) dealt with the prediction of college protest from high school protest activities. His prediction equations were best for protests over military policy (R^2 = .18) and worst for protests over administrative policy (R^2 = .08). From the current data, we may conclude that the differential predictability may be caused, at least in part, by differential response stability rather than by random sampling variability in the regressions.

The Accuracy and Precision in Self-Reports of Grades and Age[9]

Research on college students frequently relies heavily on self-reports of grades and age. The research literature typically suggests that, for continuous scales and relatively stable conditions, the overall reliability of such self-reports is in the 80-100 percent range. In many analyses, however, students are classified into strata or matched on the basis of the self-reports, in which case the reliability of the specific classification, rather than overall reliability, is most crucial. To investigate the accuracy and precision of such classification, additional analyses have been conducted on longitudinal survey data and registrar records from the Council's CIRP.

The data for these analyses were obtained from a different sample than that used in the reliability study reported in the two preceding sections. The sample consisted of 4,415 students who enrolled as freshmen from a stratified sample of 246 two-year and four-year colleges and universities in the fall of 1966, and who also participated in a mailed followup survey in the summer of 1967. College registrars supplied additional data on these same students.

Self-Reports of Freshman Grades

During the summer following their freshman year, students reported their grade-point averages by checking one of eight alternatives: D, C- and C, C+, B-, B, B+, A-, A and A+. The grades recorded on the registrars' reports were classified into the same eight levels.

To determine the nature of error in reporting, two criteria were considered: the accuracy of reporting within each grade stratum and the

---

[9]The material in this section was presented at the annual meeting of the National Council on Measurement in Education, February 6, 1971, New York.

correlation between the students' reports of their grades and the registrars' reports of the same grades. The first criterion reflects the extent to which students bias reports systematically in one direction rather than another. The second reflects the agreement between self-reports and records, as well as any unsystematic variability in self-reports.

Table 22 shows the proportions of men and women reporting grades in each of the eight strata and the proportions of those at each grade level, according to registrars' records. The biases were small in absolute terms but notable when compared with the actual percentage of students in a given cell. Men were more likely than women to report higher grades than they actually received, and this tendency was particularly evident for the A- and B+ strata. Although there were biases in the A/A+ category, the same proportions of men and women tended to be too modest in their self-reports of high grades. Among women, systematic bias in reporting grades was negligible at some grade levels, notably in the C range.

The sample included a small proportion of black students; data for 64 black men and 86 black women are presented in Table 23. Although there is some indication of greater bias for blacks of both sexes at the lower grade categories, these data were subject to much larger sampling errors so that none of the observed biases were statistically significant.

Relations between self-reported grades and those recorded in administrative records are presented in Table 24, separately by sex for the total sample and for the subsample of black students. These relations are expressed in correlational form as the ratio of Phi to Phi_{max}, thus providing an index of the veracity of self-reports relative to the maximum value Phi can have (Guilford, 1950). The overall correlation between actual grades and reported grades, not shown in the table, was .88, and was the same for men and women. The reports of students who actually performed at the A/A+ level were most likely to be accurate, and the reports of those in the B, B-, and C+ strata were least reliable. Generally, men were somewhat more likely than women to give systematically biased reports of their grades, while women were more likely to give unsystematically biased reports. The data for black students are based on such small numbers within grade categories that the statistics are too unstable for any fine interpretation.

To summarize, biases in self-reports of freshman grades depend on the grade average actually achieved and, to some extent, on the sex of the student.

Research on a much larger sample of black students is needed to verify the racial differences in reporting grades, suggested by the data on the small sample of blacks in this study. The overall veracity of self-reported grades (.88) is sufficiently high to justify their use as a proxy measure when actual grades are not available.

Self-Reports of Age

The students reported their birth dates on the 1966 SIF. In response to a separate inquiry, they gave their ages "as of December 31, 1966." Research with twins' reports of their birth dates indicates that birth dates constitute a true score, if one deletes the few who inadvertently give the current year rather than the year of birth; thus, the veracity of self-reported age can be assessed. To compare reported age with "actual" age, the date of birth was transformed to decimal form and subtracted from the decimal equivalent of December 31, 1966. Seven levels of age were considered, from 16 years or younger to 22 years or older.

Bias in self-reports of age is shown for each age level, separately by sex, in Table 25. It is evident that systematic differences between reported and computed age were insignificant. Random variability was found, however, as is reflected in the imperfect ratios of Phi to Phi_{max} presented in Table 26. There were no systematic differences found between black and nonblack students for these statistics on age.

Reported age was generally congruent with true age at all levels except the oldest. Women of 22 years or more were typically more random in their responses than were men of the same age level. The overall correlation between reported age and date of birth was somewhat higher for men (.91) than that for women (.85). Reported age is, therefore, a satisfactory proxy measure for actual age and does not require the conversion for correlational analysis that must be made when the information is obtained as "date of birth."

To summarize, self-reports of age are sufficiently accurate for matching and stratification purposes. No biases in proportions of students reporting a particular age were substantial. The extent of agreement between self-reported age and age computed from birth date was high, the only exception being that women of 22 years or older were somewhat likely to check younger age categories.

The researcher carrying out local studies can use the statistics presented here to estimate the true proportion of students who accurately report

their own age and grades. Cochran (1968) has presented equations for correcting sample proportions and these corrections can easily be generalized to 2x2 cross-tabulations and analysis. Cochran's equations require estimates of the proportion of false positives (i.e., erroneous affirmations) for a given self-report (e.g., an A+) and the proportion of false negatives (i.e., the proportion of students who actually made A+ averages, but reported a lower average); the statistics provided in Tables 22-26 should be sufficient for some of these estimates.

Errors in Initial Data Files: Verified Keypunching Vs. Optical Scanning

Even when item responses adequately represent "true scores," by meeting all the requirements of consistency, veracity, and stability, error may creep into the process of converting response marks on the measuring instrument to machine-processible form. Such error may be either random or nonrandom. To the extent that it is random, its effects on inferences drawn from analyzed data are indistinguishable from random instability of response. Thus, random error in data processing is part of the reliability problem, even though it arises from the behavior of the investigator (and his staff), rather than from that of the survey respondents.

Verified keypunching, while not absolutely error-free, introduces little error into the data actually processed, and punched cards can normally be read into a computer either directly or onto magnetic tape with negligible error rates.

With questionnaires cast into objective format, the processing of data from large-scale surveys can be more efficiently handled by optical scanning devices, which sense a respondent's answer marks and convert them into coded bit configurations on magnetic tape. Most of the survey data obtained in the Council's CIRP are so processed. The hazards involved in optical scanning include programming errors, nonreadable responses, and variations in the alignment of questionnaires being fed through the scanner. Programming errors usually result in systematic errors in the data but can be virtually eliminated with due care. Nonreadable questionnaires which result in loss of information, either on individual items or on entire respondent records can be reduced to negligible proportions if proper instructions are given to the respondent and if the administration of the forms is adequately proctored.

The maintenance of consistent alignment of the questionnaire response positions with the corresponding photoelectric cells programmed to sense specific responses is,critical for accurate data processing. Frequent checking by the operator is required to detect nonalignment. Verification by rescanning and programmed diagnostics on output from the scanner-computer complex are also useful for detecting such problems.

Usually, some coding of the responses is required. Unless the questionnaire is self-coding or precoded, coding normally precedes keypunching. In optical scanning, the coding is programmed in accordance with user specifications. In either case, coding should be exhaustive, and due allowance must be made for multiple responses and nonresponses. Otherwise, both random and nonrandom errors can be introduced into the observed scores.

Most studies of measurement error do not distinguish between that term and initial processing error, confounding the two on the assumption that the latter is either negligible or random. When this assumption is false, reported reliabilities do not have their classical meaning. If processing errors are random, it is pragmatic to combine them with random response variations when "correcting for attenuation." Nevertheless, processing error is probably more readily reduced, if not entirely eliminated, than is response variability, and indeed, there is little excuse for more than rare, sporadic errors in data processing.

Verified keypunching and optical scanning were compared using the 1969 SIF data collected during the initial testing of 191 of the subjects who participated in the study of the reliability of survey items. All questionnaires were first keypunched for use in the analysis of retest reliability: coding was accomplished directly at the keypunch, since most items are essentially precoded in a systematic way for each item format. Then the 1969 SIF questionnaires used in the pretest were optically scanned, using the same coding specifications as those used during the 1969 fall survey of entering freshmen (CIRP).

Correlations between item scores obtained from the two methods of processing are presented in Table 27. As expected, they are very high but they do not indicate perfect agreement. A spot-check against original protocols revealed no errors from the well-debugged optical scanning procedure. Practically all the discrepancies were attributable to the keypunching operations. When format changed as a new group of items were punched, the discrepancy rate

tended to rise, and then to settle back into a pattern of small and occasional discrepancies, suggesting that many errors resulted not from the keypunching per se but from "coding at the keypunch." Highest discrepancy rates (up to 10 percent) were observed on the items asking for descriptions of the parents. On many items, however, there were no discrepancies, and the rates were of the order of 1 or 2 percent. Most of the keypunch errors were confined to adjacent categories and seemed to be nonsystematic.

As a by-product of this study, some idea was obtained of the rate of missing data, including nonprocessible responses and omitted items, which never exceeded 5 percent of the subjects for a given item, and, most typically, was of the order of 1 percent. Data were missing more commonly in the case of the more sensitive items (e.g., income) or where the information sought might not be known or meaningful to the respondent (e.g., reporting about parents, stating expectations, having an opinion on certain issues).

The general conclusion is that errors of measurement arise primarily from sources other than those associated with the initial coding and transcribing of item responses to machine-processible form, provided that reasonable precautions and checks are incorporated into such operations.

Concluding Remarks

The major reason for concern about the reliability of survey measurements is that random errors in data may lead to false or misleading inferences. These random errors may occur either as response instability (error of measurement) or as errors in processing the data. In either case, inferential error may result from errors in the classification of observation units or from distortions in a correlation matrix by differential attenuation of observed correlations. Although these principles are widely known, especially with respect to psychological tests and measurements, they are frequently ignored in social and educational research based on survey questionnaire procedures for collecting data. This report has presented empirical estimates of reliability for the kinds of variables typically used in the annual surveys of entering college freshmen conducted by the American Council on Education.

Despite some limitations in the sample studied, these estimates are satisfactory for use in analyzing the survey data. The estimates may also be useful when analyzing panel data obtained by posttesting with the same variables in a followup study, and when independent estimates of the reliabilities of the followup measurements are not available.

Except for grossly awkward or otherwise inappropriate administrative conditions, the major factors determining reliability are the content and structure of an item. For this reason, there may be some justification for relying on some of the reported reliability data in other surveys; wherever possible, however, independent estimates of the reliabilities of survey variables should be obtained. The substantive content of an item must be relevant and meaningful to the entire heterogeneous population of respondents, a condition that requires a compromise between sufficient generality to ensure that heterogeneity is represented in the response possibilities, and sufficient clarity to ensure that any given response will be interpretable. Moreover, the various response alternatives in a multiple-choice item must be both exhaustive and mutually exclusive. The semantic content of an item is critical for both veracity and stability of response: not only must the vocabulary be basic enough for all respondents to understand what is being asked (note, e.g., the difficulties some freshmen have with a term like Caucasian), but individual terms must carry a common meaning to all the subpopulations being queried.

Generally, an increase in the number of meaningful response alternatives in a multiple-choice item increases reliability and results in more homogeneous interpretability of the responses to any one alternative. In the case of ordered categories, however, a point of diminishing returns may be reached rather quickly as the respondent is required to make finer discriminations. In a checklist item in which each subitem results in a separate dichotomous variable, considerable variation in reliabilities may be expected as a function of the base rates (popularities) of various subitems and of the probability that subjects will respond to more than one subitem in the list.

Reliabilities of items also vary as a function of their content. Consistently, in this and in previous studies, objective items asking for information about the respondent, especially demographic and family background information, had high reliabilities as did items asking for academic and career plans and aspirations. This is fortunate in that many analyses of survey data involve breakouts in terms of the categories in such items. In addition, items asking the respondent to recall verifiable facts about his recent activities and achievements generally have high reliabilities. Lower reliabilities are typical of items asking the respondent for his attitudes and opinions or for his perceptions of his parents' characteristics. Generally lower reliabilities are also associated with items that depend upon the behavior of others and for

items oriented toward future events (e.g., goals and expectations), though in the latter case, reliability varied considerably for various specific future events.

With rare exceptions, the reliability data obtained in this study indicate that all of the variables considered are sufficiently reliable for analytic use in statistical research, and many of them are sufficiently reliable that inferences drawn from analyses taking measurement errors into account can be used in institutional decision making and in counseling students.

Bibliography

The ACE Office of Research: Its Purposes and Activities. Prepared by L. Kent with the cooperation of the staff of the Office of Research. Washington: American Council on Education, 1972, forthcoming.

American Psychological Association. Technical Recommendations for Psychological Tests and Diagnostic Techniques. Washington: The Association, 1954.

Assukal, K., and Procter, C.H. Testing Hypotheses with Categorical Data Subject to Misclassification. Mimeograph Series No. 448, Institute of Statistics, University of North Carolina, September 1965.

Astin, A.W. "A Re-Examination of College Productivity." Journal of Educational Psychology, 1961, 52, 173-78.

Astin, A.W. "'Productivity' of Undergraduate Institutions." Science, 1962, 136, 129-35.

Astin, A.W. "Personal and Environmental Factors Associated with College Dropouts Among High Aptitude Students." Journal of Educational Psychology, 1964, 55, 219-27.

Astin, A.W. Who Goes Where to College? Chicago: Science Research Associates, 1965.

Astin, A.W. "Personal and Environmental Determinants of Student Activism." Measurement and Evaluation in Guidance, 1968, 1, 149-61.

Astin, A.W. "The Methodology of Research on College Impact." Sociology of Education, 1970, 43, 223-52; 347-50.

Astin, A.W. "New Evidence on Campus Unrest, 1969-70." Educational Record, 1971, 52, 41-46.

Astin, A.W., and Panos, R.J. The Educational and Vocational Development of College Students. Washington: American Council on Education, 1969.

Astin, A.W., Panos, R.J., and Creager, J.A. A Program of Longitudinal Research on the Higher Educational System. ACE Research Reports, Vol. 1, No. 1, 1966.

Bailar, B.A. "Recent Research in Reinterview Procedure." Journal of the American Statistical Association, 1968, 63, 41-63.

Bailey, D., and Schotta, C. "The Private Rate of Return to Investment in Undergraduate Education." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1969.

Bayer, Alan E. "Construction of a Race Item for Survey Research." Mimeographed. Washington: American Council on Education, 1971.

Bibliography (Con't.)

Bayer, A.E., Astin, A.W., and Boruch, R.F. "College Students' Attitudes Toward Social Issues, 1967-70." Educational Record, 1971, 52, 52-59.

Bayer, A.E., Astin, A.W., Boruch, R.F., and Creager, J.A. Users' Manual: ACE Higher Education Data Bank. ACE Research Reports, Vol. 4, No. 1, 1969.

Bayer, A.E., Drew, D.E., Astin, A.W., Boruch, R.F., and Creager, J.A. The First Year of College: A Follow-up Normative Report. ACE Research Reports, Vol. 5, No. 1, 1970.

Boruch, R.F. "Computation of Ø and Ømax when the Data are Subject to Misclassification." Educ. and Psych. Measurement, 1972, in press.

Boruch, R.F., and Creager, J.A. "A Note on Stability of Self-Reported Protest Activity and Attitudes." Measurement and Evaluation in Guidance, 1972, 5, in press.

Boruch, R.F., and Dutton, J.E. "A Program for Testing Hypotheses About Correlation Arrays." Educ. and Psych. Measurement, 1970, 30, 719-21.

Bradburn, N.M. "Selecting the Questions to Be Asked." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1969, 178-81.

Campbell, D.T. "The Effects of Colleges on Students: Proposing a Quasi-experimental Design Approach." Mimeographed. Evanston, Ill.: Department of Psychology, Northwestern University, 1967.

Campbell, D.T. "Quasi-experimental Design." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1965, 157-60.

Campbell, D.T., and Fiske, D.W. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." Psychological Bulletin, 1959, 56, 81-105.

Campbell, D.T., Siegman, C.R., and Rees, M.B. "Direction of Wording Effects in the Relationships Between Scales." Psychological Bulletin, 1967, 8, 293-303.

Campbell, D.T., and Stanley, J.C. Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally, 1966.

Cochran, W.G. "Errors of Measurement in Statistics." Technometrics, 1968, 10, 637-65.

Bibliography (Con't.)

Creager, J.A., Astin, A.W., Boruch, R.F., Bayer, A.E., and Drew, D.E. National Norms for Entering College Freshmen -- Fall 1969. ACE Research Reports, Vol. 4, No. 7, 1969.

Cronbach, L.J. Essentials of Psychological Testing. New York: Harper and Row, 1960.

DeGracie, J.D. "Analysis of Covariance when the Concomitant Variable Is Measured with Error." Ph.D. dissertation, Iowa State University, 1968.

Dillon, D. Memorandum on accuracy of mark-sensing equipment. Mimeographed. Minneapolis, National Computer Systems, 1970.

DuBois, N.S. "A Solution to the Problem of Linking Multivariate Documents." Journal of the American Statistical Association, 1969, 64, 163-74.

Dugan, D.J. "The Impact of Parental and Educational Investments." Paper presented at the annual meeting of the American Statistical Association, New York City, August 1969.

Edwards, A.L. Techniques of Attitude Scale Construction. New York: Appleton-Century-Crofts, 1957.

Edwards, K.J. "Correcting Partial, Multiple, and Canonical Correlations for Attenuations." Paper presented at the annual meeting of the American Educational Research Association, New York City, 5 February 1971.

Finucci, J.M. "Early History and Basic Derivation of Attenuation Formulas." Paper presented at the annual meeting of the American Educational Research Association, New York City, 5 February 1971.

Freund, R.F., and Hartley, H.O. "A Procedure for Automatic Data Editing." Journal of the American Statistical Association, 1967, 62, 341-52.

Giesbrect, F. "Classification Errors and Measures of Association in Contingency Tables." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1967.

Gleser, L.J. "The Attenuation Paradox and Internal Consistency." Paper presented at the annual meeting of the American Educational Research Association, New York City, 5 February 1971.

Greenberg, B.G., Abul-Ela, A.A., Simmons, W.R., and Horvitz, D.G. "The Unrelated Question Randomized Response Model: Theoretical Framework." Journal of the American Statistical Association, 1969, 64, 520-39.

Guilford, J.P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Co., 1950.

Bibliography (Con't.)

Guthrie, H.M. "Some Methodological Issues in Validation Studies." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1965, 193-96.

Haber, L.D. "Evaluating Response Error in the Reporting of Income of the Aged: Benefit Income." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1966.

Hansen, M.H., Hurwitz, W.N., and Bershad, M.A. "Measurement Errors in Censuses and Surveys." Bulletin of the International Statistical Institute, 1961, 38, 359-74.

Hansen, M.H., Hurwitz, W.N., and Pritzker, L. "Estimation and Prediction of Gross Differences and Simple Response Variance." Contributions to Statistics Presented to Professor P.C. Mahalanobis. Oxford, England: Pergsmon Press, 1964.

Harman, H.H. Modern Factor Analysis. Chicago: University of Chicago Press, 1960.

Harris, C.W. (Ed.) Problems in Measuring Change. Madison: University of Wisconsin Press, 1967.

Johnston, J. Econometric Methods. New York: McGraw-Hill Book Co., 1963.

Joreskog, K.G., and Gruvaeus, G. RMLFA: A Computer Program for Restricted Maximum Likelihood Factor Analysis. Research Bulletin RB-67-21. Princeton, N.J.: Educational Testing Service, 1967.

Keating, E., Peterson, D.G., and Stone, C.H. "Validity of Work Histories Obtained by Interview." Journal of Applied Psychology, 1950, 34, 6-11.

Kish, L. Survey Sampling. New York: John Wiley and Sons, 1967.

Koch, G.G. "The Effect of Nonsampling Errors on Measures of Association in 2 x 2 Contingency Tables." Journal of the American Statistical Association, 1969, 64, 851-63.

Komorita, S.S., and Graham, W.K. "Number of Scale Points and the Reliability of Scales." Educational and Psychological Measurement, 1965, 4, 987-95.

Livingston, S.A., and Stanley, J.C. What Is the True Coefficient of Partial Correlation? Center for Social Organization of the Schools Report No. 107. Baltimore: Johns Hopkins University, July 1971.

Lord, F.M., and Novick, M.R. Statistical Theories of Mental Test Scores. Reading, Mass.: Addison Wesley Publishing Co., 1968.

Madansky, A. "The Fitting of Straight Lines when Both Variables Are Subject to Error." Journal of the American Statistical Association, 1959, 54, 173-205.

Bibliography (Con't.)

Madow, W.G. "On Some Aspects of Response Error Measurement." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1965, 182-92.

Maxey, E.J., and Ormsby, V.J. The Accuracy of Self-Report Information on the ACT Test Battery: High School Grades and Items of Nonacademic Achievement. ACT Research Report No. 45. Iowa City: American College Testing Program, July 1971.

Maynes, E.S. "Minimizing Response Errors in Financial Data: The Possibilities." Journal of the American Statistical Association, 1968, 63, 214-27.

Menne, J.W. Klingensmith, J.E., and Nord, D. "The Problem of Measurement Accuracy in Educational Research." Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February, 1969.

Minton, G. "Inspection and Correction Error in Data Processing." Journal of the American Statistical Association, 1969, 64, 1256-75.

Mosteller, F. "Association and Estimation in Contingency Tables." Journal of the American Statistical Association, 1968, 63, 1-27.

O'Reagan, R.T. "Relative Costs of Computerized Error Inspection Plans." Journal of the American Statistical Association, 1969, 64, 1245-56.

Parten, M. Surveys, Polls, and Samples. New York: Harper, 1950.

Peterson, R.E. Technical Manual: College Student Questionnaire. Princeton, N.J.: Institutional Research Program for Higher Education, 1965.

Porter, A.C. "The Effects of Using Fallible Variables in the Analysis of Covariance." Ph.D. dissertation, University of Wisconsin, 1967.

Porter, A.C. "How Errors of Measurement Affect ANOVA, Regression Analyses, ANCOVA, and Factor Analyses." Paper presented at the annual meeting of the American Educational Research Association, New York City, 5 February 1971.

Press, S.J. "Estimating from Misclassified Data." Journal of the American Statistical Association, 1968, 63, 123-33.

Schotta, C., and Hoffman, R.B. "Some Statistical Aspects of Educational Evaluation." Proceedings of the Social Statistics Section. Washington: American Statistical Association, 1968.

Siegel, P.M., and Hodge, R.W. "A Causal Approach to the Study of Measurement Error." In H.M. Blalock and A.B. Blalock (Eds.), Methodology in Social Research. New York: McGraw-Hill Book Co., 1968. Pp. 28-59.

Bibliography (Con't.)

Smith, N. "A Note on Stability in Questionnaire Response." American Journal of Sociology, 1932, 37, 713-20.

Snedecor, G.W., and Cochran, W.G. Statistical Methods. Ames: Iowa State University Press, 1967.

Spencer, R.C. "Reliability and Validity of the Digitek Optical Scanner in Test Scoring Operations." Educational and Psychological Measurement, 1966, 26, 719-23.

Stanley, J.C., and Livingston, S.A. "Correcting Eta Squared, Epsilon Squared, Omega Squared, and Intraclass r for Attenuation." Paper presented at the annual meeting of the American Educational Research Association, New York City, 5 February 1971.

Stephan, F.F. "The Quality of Statistical Information and Statistical Inference in a Rapidly Changing World." Journal of the American Statistical Association, 1967, 62, 1-9.

Thurstone, L.L. Vectors of the Mind. Chicago: University of Chicago Press, 1935.

Tucker, L.R., Damarin, F., and Messick, S. "A Base-Free Measure of Change." Psychometrika, 1966, 4, 457-74.

Valentine, L.D., Jr. "Rejection Rates for 100 Mark-Sense Cards Processed Through the IBM 519." Memorandum WCLL-TM-59-29. Lackland Air Force Base, Texas, 1959.

Valentine, L.D., Jr. "Effects of Marking Instructions on Rejection Rates of Mark-Sense Cards Processed Through the IBM 519." Memorandum WRDP-TM-60-22. Lackland Air Force Base, Texas, 1960.

Walsh, W.B. "Validity of the Self-Report." Journal of Counseling Psychology, 1967, 14, 18-23.

Walsh, W.B. "Validity of the Self-Report: Another Look." Journal of Counseling Psychology, 1968, 15, 180-86.

Wolins, L. "The Use of Multiple Regression Procedures when the Predictor Variables Are Psychological Tests." Educational and Psychological Measurement, 1967, 27, 821-27.

Wolins, L. "Some Computational, Methodological and Theoretical Considerations Relevant to Summarizing Psychological Data." Mimeographed. Ames: Iowa State University, 1968.

Tables

Table 1

Test-Retest Statistics and Reliabilities for Items on Demographic Characteristics, Family Background, High School Achievements, and College Admissions

Item	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r
Age	4.23	.89	4.24	.94	.96
Father's education	4.12	1.41	4.12	1.41	.99
Mother's education	3.65	1.08	3.63	1.08	.97
Estimated parents' income	6.11	1.89	6.15	1.88	.98
Concern about financing college education	1.65	.62	1.56	.58	.85
Average high school grade	4.90	1.64	4.88	1.64	.98
Rank in high school class	3.24	3.46	3.34	3.53	.97
Number of applications to other colleges	3.02	1.72	3.01	1.75	.98
Number of acceptances by other colleges	2.35	1.33	2.32	1.34	.97
Distance of college from home	2.72	3.03	2.69	3.02	.97

Table 2

Bivariate Distribution for Stability of Reported Father's Education[a]

	Father's Education	1	2	3	4	5	6	Row Percentage
1	Grammar school or less	3.5						3.5
2	Some high school		8.9	1.0				9.9
3	High school graduate		1.0	21.8				22.8
4	Some college				18.8	1.0		19.8
5	College degree					23.2		23.2
6	Postgraduate degree					1.0	19.8	20.8
	Column Percentage	3.5	9.9	22.8	18.8	25.2	19.8	100.0

[a]In these and all subsequent bivariate tables, rows refer to test data and columns refer to retest data.

Table 3

Bivariate Distribution for Stability of Reported Mother's Education[a]

	Mother's Education	1	2	3	4	5	6	Row Percentage
1	Grammar school or less	1.0						1.0
2	Some high school		7.0	0.5				7.5
3	High school graduate		1.0	46.0				47.0
4	Some college			0.5	20.0	0.5		21.0
5	College degree			0.5	1.0	15.0	0.5	17.0
6	Postgraduate degree					0.5	6.0	6.5
	Column Percentage	1.0	8.0	47.5	21.0	16.0	6.5	100.0

Table 4

Bivariate Distribution for Stability of Estimated Parents' Income

	Estimated Parents' Income	1	2	3	4	5	6	7	8	9	Row Percentage
1	Less than $4,000	1.0									1.0
2	$4,000 - $5,999	0.5	3.0								3.5
3	$6,000 - $7,999		0.5	3.0	0.5						4.0
4	$8,000 - $9,999				4.5	1.5					6.0
5	$10,000 - $14,999					21.5	3.0		0.5		25.0
6	$15,000 - $19,999					0.5	21.5	1.0			23.0
7	$20,000 - $24,999						0.5	11.0	0.5		12.0
8	$25,000 - $29,999							1.5	8.5		10.0
9	$30,000 Or more									15.5	15.5
	Column Percentage	1.5	3.5	3.0	5.0	23.5	25.0	13.5	9.5	15.5	100.0

Table 5

Bivariate Distribution for Stability of Reported Concern About Financing College Education

	Concern About Financing Education	1	2	3	Row Percentage
1	None	40.3	2.0		42.3
2	Some	6.5	43.3		49.8
3	Major		3.5	4.5	8.0
	Column Percentage	46.8	48.8	4.5	100.0

Table 6

Bivariate Distribution for Stability of Reported Average High School Grade

	Average High School Grade	1	2	3	4	5	6	7	8	Row Percentage
1	A or A+	4.0	0.5							4.5
2	A-	0.5	9.9	0.5						10.9
3	B+		1.0	16.8	2.5					20.3
4	B			1.5	28.2	2.0				31.7
5	B-				1.0	11.9	0.5			13.4
6	C+						7.9	0.5		8.4
7	C						1.0	6.4	0.5	7.9
8	D							0.5	2.5	3.0
	Column Percentage	4.5	11.4	18.8	31.7	13.9	9.4	7.4	3.0	100.0

Table 7

Bivariate Distribution for Stability of Reported Rank in High School Class

	Rank in High School Class	1	2	3	4	5	6	Row Percentage
1	Top 1 percent	1.6		0.5			0.5	2.6
2	Top 10 percent		25.1	0.5	1.1			26.7
3	Top Quarter	0.5		28.3	1.6			30.1
4	Second Quarter		0.5	1.1	20.3	0.5		21.4
5	Third Quarter				1.1	12.3		13.4
6	Fourth Quarter					0.5	3.7	4.2
	Column Percentage	2.1	25.7	30.5	24.1	13.4	4.3	100.0

Table 8

Bivariate Distribution for Stability of Reported Distance of College from Home

	Distance of College From Home	1	2	3	4	5	6	Row Percentage
1	10 or fewer miles	24.6						24.6
2	11-50 miles	1.6	27.3	1.1		0.5		30.5
3	51-100 miles			4.9				4.9
4	101-500 miles			0.5	32.8			33.3
5	501-1,000 miles				0.5	2.7		3.2
6	1,000 or more miles				0.5		2.7	3.2
	Column Percentage	26.2	27.3	6.5	33.8	3.2	2.7	99.7

Table 9

Bivariate Distribution for Stability of Reported Region of Birth[a]

	Region of Birth	1	2	3	4	5	6	7	Row Percentage
1	Midwest	69.5				0.5			70.0
2	New England		2.0						2.0
3	North-Central			8.5					8.5
4	Northwest				0.5				0.5
5	Southern	0.5				12.5			13.5
6	Western						3.0		3.0
7	Foreign Country	0.5						2.0	2.5
	Column Percentage	70.5	2.0	8.5	0.5	13.0	3.0	2.0	99.5

[a]98 percent concordance.

Table 10

Test-Retest Statistics and Reliabilities
for Four Reported Sources of Financial Support During Freshman Year[a]

Source of Financial Support During Freshman Year	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	$\underline{r}$
Personal savings or employment	1.82	.69	1.90	.68	.86
Aid from parents or family	2.52	.77	2.52	.74	.88
Repayable loan	1.23	.58	1.26	.60	.86
Scholarship/grant/other gift	1.41	.73	1.37	.69	.85

[a]For each source, the respondent was asked to check whether it was major (coded 3), minor (coded 2), or not a source (coded 1).

Table 11

Test-Retest Statistics and Reliabilities
for Reported Attributes of Parents

Parents' Attributes	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	$\underline{r}$
Interested in intellectual pursuits	3.11	.81	3.09	.75	.71
Interested in cultural pursuits	2.94	.75	2.88	.72	.73
Deeply religious	2.40	.82	2.35	.81	.82
Interested in politics	2.95	.64	2.95	.60	.65
Deeply concerned about children	3.74	.59	3.75	.57	.73
Financially comfortable	3.21	.75	3.15	.68	.73
Have high aspirations for respondent	3.64	.61	3.64	.55	.60

Table 12

Test-Retest Statistics and Reliabilities
for Reported Chances of Future Events

Future Event[a]	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r
Getting married while in college	2.26	.99	2.27	.97	.88
Marrying within a year after college	2.84	1.00	2.84	.98	.82
Obtaining average grade of A- or higher	2.12	.85	2.14	.87	.77
Changing major field	2.84	.93	2.77	.95	.81
Changing career choice	2.82	.99	2.74	.97	.80
Failing one or more courses	2.31	.93	2.26	.89	.76
Graduating with honors	2.29	.88	2.23	.85	.73
Being elected to a student office	1.91	.81	1.98	.79	.73
Joining a social fraternity or sorority	2.39	1.17	2.34	1.11	.86
Authoring a published article	2.10	.94	2.04	.90	.76
Being drafted while in college	1.41	.75	1.48	.75	.80
Being elected to an honor society	2.17	.92	2.16	.89	.77
Protesting over U.S. military policy	2.70	1.15	2.63	1.10	.88
Protesting over college administrative policy	2.59	1.01	2.54	.96	.84
Protesting over racial/ethnic policy	2.72	1.07	2.61	1.04	.83
Dropping out temporarily	2.05	.86	1.95	.82	.69
Enlisting in armed services before graduation	1.18	.51	1.27	.61	.62
Being more successful than average	3.04	.61	3.07	.60	.59
Dropping out permanently	1.40	.69	1.45	.69	.58
Transfering to another college	2.67	1.00	2.62	.99	.82

[a]Alternate responses and scoring key: Very good chance = 4; some chance = 3; very little chance = 2; no chance = 1.

Table 13

Test-Retest Statistics and Reliabilities
for Reported Objectives

Objective[a]	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r
Being accomplished in a performing art	1.63	.84	1.78	.90	.78
Being an authority in field	2.82	.81	2.85	.87	.73
Obtaining recognition from peers	2.41	.90	2.47	.88	.68
Influencing the political structure	2.01	.87	2.03	.86	.72
Influencing social values	2.41	.92	2.43	.87	.71
Raising a family	3.08	.98	3.16	.96	.87
Having an active social life	2.74	.89	2.75	.88	.74
Having friends different from self	2.80	.89	2.85	.85	.70
Being an expert in finance and commerce	1.56	.77	1.63	.83	.74
Having administrative responsibility for work of others	1.90	.86	1.98	.87	.66
Being very well-off financially	2.45	.85	2.47	.80	.81
Helping others in difficulty	2.92	.79	2.84	.81	.65
Becoming a community leader	1.83	.85	1.91	.81	.74
Contributing to a scientific theory	1.37	.67	1.38	.73	.79
Writing original works	1.76	.92	1.80	.98	.80
Not being obligated to people	2.08	1.03	2.12	1.00	.71
Creating works of art	1.86	.99	1.89	.94	.81
Keeping up with political affairs	2.74	.83	2.65	.88	.81
Succeeding in own business	2.05	1.04	2.16	1.03	.67
Developing a philosophy of life	3.35	.83	3.35	.79	.69

[a]Alternatives and scoring key: Essential = 4; very important = 3; somewhat important = 2; not important = 1.

Table 14

Test-Retest Statistics and Reliabilities
for Reported Attitudes Toward Federal Involvement in Problem Areas[a]

Problem Area	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r
Control of cigarette advertising	3.48	1.13	3.46	1.10	.73
Elimination of violence from TV	2.78	1.17	2.94	1.15	.64
Control of pollution	4.71	.59	4.64	.64	.43
Control of birth rate through tax incentives	3.22	1.37	3.27	1.18	.63
Consumer protection	4.14	.73	4.09	.67	.41
Compensatory education for the disadvantaged	3.98	.79	3.70	.82	.68
Special benefits for veterans	3.28	.74	3.19	.72	.58
Control of firearms	3.85	1.05	3.75	1.04	.79
Elimination of poverty	4.39	.85	4.27	.82	.69
Crime prevention	4.49	.73	4.36	.69	.44
School desegregation	3.83	1.24	3.75	1.13	.83
Financial aid for disadvantaged	3.67	.88	3.50	.83	.57
Control of student activists	2.55	1.22	2.53	1.12	.69

[a]Alternatives and scoring key: Initiate new crash programs = 5; Increase involvement from current level = 4; Maintain current level of involvement = 3; Decrease involvement from current levels = 2; Eliminate any existing programs or remain uninvolved = 1.

Table 15

Test-Retest Statistics and Reliabilities
for Reported Attitudes about Campus and Social Issues[a]

Item	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r
Students should help design curriculum	3.36	.76	3.28	.71	.64
Scientists should publish all findings	2.75	.93	2.72	.88	.63
Individual cannot change society	2.20	.94	2.25	.84	.62
Colleges have right to control behavior of students off campus	1.22	.59	1.27	.60	.48
Chief benefit of college is monetary	2.17	.96	2.31	.92	.72
Faculty promotions should be based on student evaluations	2.87	.86	2.86	.80	.57
My beliefs are similar to those of other students	2.58	.69	2.60	.68	.66
College officials should clear student publications	1.93	.90	1.81	.78	.59
Marijuana should be legalized	2.76	1.10	2.75	1.08	.88
College has right to ban extremist speakers	1.65	.87	1.73	.88	.61
Army should be voluntary	2.92	.98	2.88	.94	.69
Disadvantaged should be given preferential treatment in admissions	2.20	.88	2.25	.90	.74
College officials too lax with student protests	2.06	.85	2.12	.86	.66

[a]Alternative and scoring key: Agree strongly = 4, agree somewhat = 3, disagree somewhat = 2, disagree strongly = 1.

Table 16

Test-Retest Response Probabilities and Phi Coefficients for Checklist of High School Achievements

High School Achievement	P_1	P_2	Ø
Elected president of student organization(s)	.26	.25	.90
Received high rating in state/regional music contest	.10	.08	.89
Participated in state/regional speech/debate contest	.07	.07	.92
Had major part in play	.21	.20	.96
Won varsity letter (sports)	.28	.30	.96
Won award in art competition	.07	.06	.88
Edited school paper, yearbook, literary magazine	.16	.17	.91
Had original writing published	.23	.26	.88
Was member of scholastic honor society	.25	.25	.96
Received National Merit recognition	.13	.13	1.00

Table 17

Stability of Reported Current Political Preference

	Current Political Preference	1	2	3	4	5	Row Percentage
1	Left	8.2	1.6	0.5	0.5		10.9
2	Liberal	2.2	45.7	2.7	1.1		51.6
3	Middle of the road		2.7	17.9	1.1		21.7
4	Moderately conservative			1.6	10.9	0.5	13.0
5	Strongly conservative				1.6	1.1	2.7
	Column Percentage	10.3	50.0	22.8	15.2	1.6	100.0

Table 18

Stability of Expected Political Preference in Four Years

	Expected Political Preference in Four Years	1	2	3	4	5	Row Percentage
1	Left	9.2			0.5		9.8
2	Liberal	1.1	45.7	2.7			49.5
3	Middle of the road		2.7	22.3			25.0
4	Moderately conservative			2.3	12.0		14.1
5	Strongly conservative				0.5	1.1	1.6
	Column Percentage	10.3	48.4	27.2	13.0	1.1	100.0

Table 19

Stability of Reported Participation in Protest Against Racial, Military, or Administrative Policy During Senior Year in High School[a]

Institution	Protest Issue	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r	$\Delta\bar{X}_{2-1}$	ΔS.D.$_{2-1}$
Public university (N = 97)								
	Race	1.31	.51	1.35	.58	.69	.04	.07
	Military	1.42	.59	1.42	.59	.76	.00	.00
	Administration	1.31	.47	1.38	.57	.36	.07	.10
Private university (N = 62)								
	Race	1.61	.64	1.72	.61	.54	.11	-.03
	Military	2.03	.66	1.93	.66	.79	-.10	.00
	Administration	1.53	.67	1.47	.62	.59	-.06	-.05
Community college (N = 43)								
	Race	1.63	.43	1.19	.50	.74	-.44*	.07
	Military	1.63	.43	1.12	.40	.93	-.51*	-.03
	Administration	1.63	.37	1.09	.29	.29	-.54*	-.08

[a]Alternatives and scoring key: Frequently =3; occasionally = 2; not at all = 1. In this table and in Tables 20 and 21, a correlation of .30 is necessary to be fairly sure ($p > .95$) that the statistic is significant with degrees of freedom of at least 40.

*Difference significant at the .05 level.

Table 20

Stability of Freshmen Predicting Future Involvement in Protests about Racial, Military, and Administrative Policy[a]

Institution	Protest Issue	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r	$\Delta\bar{X}_{2-1}$	ΔS.D.$_{2-1}$
Public university (N = 97)								
	Race	2.67	1.04	2.51	1.00	.73	-.16	-.04
	Military	2.62	1.08	2.54	1.00	.85	-.12	-.08
	Administration	2.67	.93	2.56	.92	.82	-.11	-.01
Private university (N = 62)								
	Race	3.26	.84	3.18	.84	.81	-.08	.00
	Military	3.35	.87	3.27	.87	.86	-.08	.00
	Administration	3.11	.81	3.09	.72	.74	-.02	-.09
Community college (N = 43)								
	Race	2.07	1.07	2.04	1.00	.80	-.03	-.07
	Military	1.95	1.15	1.91	1.07	.83	-.04	-.08
	Administration	1.65	.84	1.72	.73	.73	.07	-.11

[a]Alternatives and scoring key: Very good chance = 4; some chance = 3; very little chance = 2; no chance = 1.

Table 21

Stability of Reported Attitudes on Items Pertaining to Student Freedom and Administrative Control

Institution	Item	$\bar{X}_1$	S.D.$_1$	$\bar{X}_2$	S.D.$_2$	r	$\Delta\bar{X}_{2-1}$	ΔS.D.$_{2-1}$
Public university (N = 97)	Liberalism[a]	3.34	.87	3.43	.86	.93	.09	-.01
	Control activists[b]	2.56	1.12	2.57	1.07	.57	.01	-.05
	Regulate off-campus behavior[c]	1.25	.69	1.27	.62	.45	.02	-.24
	Regulate publications[c]	1.96	.95	1.85	.76	.62	-.09	-.21
	Ban speakers[c]	1.66	.89	1.72	.87	.49	.06	-.01
	Administrative laxity[c]	1.99	.77	2.16	.84	.70	.17	-.07
Private university (N = 62)	Liberalism[a]	3.92	.67	3.97	.64	.89	.05	-.03
	Control activists[b]	2.15	1.17	2.13	1.00	.85	-.02	-.17
	Regulate off-campus behavior[c]	1.13	.42	1.22	.52	.61	.09	.10
	Regulate publications[c]	1.62	.75	1.57	.67	.45	-.05	-.08
	Ban speakers[c]	1.44	.74	1.44	.72	.77	.00	-.02
	Administrative laxity[c]	1.94	.79	1.89	.79	.60	-.05	.00
Community college (N = 43)	Liberalism[a]	3.19	1.11	3.15	1.02	.87	-.04	-.09
	Control activists[b]	3.12	1.26	3.02	1.19	.62	-.10	-.07
	Regulate off-campus behavior[c]	1.30	.56	1.35	.65	.49	.05	.09
	Regulate publications[c]	2.28	.83	2.07	.91	.58	-.21	.08
	Ban speakers[c]	1.93	.91	2.14	.96	.63	.21	.05
	Administrative laxity[c]	2.40	.98	2.37	.95	.63	-.03	-.03

[a]Self-rating. Alternatives and scoring key: Highest 10 percent = 5; above average = 4; average = 3; below average = 2; lowest 10 percent = 1.

[b]Federal involvement. See Table 14, footnote a, for alternatives and scoring key.

[c]Attitude item. Alternatives and scoring key: Agree strongly = 4; ageee somewhat = 3; disagree somewhat = 2; disagree strongly = 1.

Table 22

Agreement Between Self-Reported and Recorded Freshman Grades
(Proportions)

		Men (N = 2,090)			Women (N = 2,325)		
Category	Range[b]	Self-Reported	Recorded	Bias[a]	Self-Reported	Recorded	Bias[a]
A/A+	3.67-4.00	.02	.03	-.01	.02	.03	-.01
A-	3.50-3.66	.04	.02	+.02	.05	.04	+.01
B+	3.16-3.49	.08	.06	+.02	.12	.11	+.01
B	2.83-3.15	.11	.10	+.01	.17	.16	+.01
B-	2.50-2.82	.16	.19	-.03	.18	.20	-.02
C+	2.16-2.49	.23	.20	-.03	.20	.20	.00
C/C-	1.50-2.15	.29	.31	-.02	.22	.22	.00
D	0-1.49	.07	.09	-.02	.03	.04	-.01
GPA	(0 - 4)	1.78	1.65	+.13	1.96	1.94	+.02

[a]Bias refers to the difference between the self-reported and recorded values, and all nonzero differences are significant at the .005 level.

[b]This type of data should be replicated for other conversions between numerical and letter grades in common use. Different numerical ranges associated with each letter grade will probably change both self-reported and recorded grade proportions but have little effect on their differences or bias.

Table 23

Agreement Between Self-Reported and Recorded Freshman Grades of Black Students

		Men (N = 64)			Women (N = 86)		
Category	Range[b]	Self-Reported	Recorded	Bias[a]	Self-Reported	Recorded	Bias[a]
A/A+	3.67-4.00	.00	.00	.00	.00	.00	.00
A-	3.50-3.66	.02	.02	.00	.00	.01	-.01
B+	3.16-3.49	.03	.02	+.01	.03	.01	+.02
B	2.83-3.15	.03	.03	.00	.10	.08	+.02
B-	2.50-2.82	.22	.22	.00	.14	.13	+.01
C+	2.16-2.49	.28	.23	+.05	.26	.28	-.02
C/C-	1.50-2.15	.38	.42	-.04	.40	.31	+.09
D	0-1.49	.05	.06	-.01	.07	.17	-.10

[a]Bias refers to the difference between the self-reported and recorded values, and all nonzero differences are significant at the .005 level.

[b]This type of data should be replicated for other conversions between numerical and letter grades in common use. Different numerical ranges associated with each letter grade will probably change both self-reported and recorded grade proportions but have little effect on their differences or bias.

Table 24

$\emptyset$ and $\emptyset/\emptyset_{max}$ for Self-Reported and Recorded Freshman Grades, by Sex and Race[a]

Group	Coefficient	Grade A/A+	A-	B+	B	B-	C+	C/C-	D
All men	$\emptyset$	.69	.49	.56	.42	.39	.40	.59	.60
	$\emptyset/\emptyset_{max}$	.70	.59	.65	.44	.42	.43	.62	.71
All women	$\emptyset$	.56	.41	.52	.43	.38	.37	.62	.46
	$\emptyset/\emptyset_{max}$	.78	.43	.56	.44	.41	.38	.63	.53
Black men (N=64)	$\emptyset$	.00	1.00	.70	1.00	.36	.39	.52	.55
	$\emptyset/\emptyset_{max}$	.00	1.00	1.00	1.00	.36	.44	.57	.64
Black women (N=86)	$\emptyset$	.00	.00	-.02	.73	.45	.41	.58	.60
	$\emptyset/\emptyset_{max}$	.00	.00	-.04	.84	.47	.43	.69	1.00

[a]In the case of dichotomous variables generated from the responses to a single category, the reliability as measured by the Phi coefficient is a function of the base rate, i.e., the frequency with which that category is checked. This, in turn, is not independent of the number of, and base rates for, the other categories. Hence, the Phi coefficients may be low even when there is a high concordance on retest. The Phi/Phi_{max} coefficients, which allow for variation in base rates, are much higher and more consistent across categories, but have the disadvantage of representing an intrinsic relationship rather than one which is practical in correcting for attenuation in correlational analysis (Guilford, 1950).

Table 25

Self-Reported and Computed Proportions of Students in Age Categories

Category	Men (N = 2,090) Self-Reported	Computed	Bias	Women (N = 2,325) Self-Reported	Computed	Bias
Age $\geq$ 22	.02	.02	0	.01	.01	0
Age = 21	.01	.01	0	.01	.01	0
Age = 20	.01	.01	0	.01	.01	0
Age = 19	.11	.10	.01	.08	.07	.01
Age = 18	.80	.80	0	.85	.85	0
Age = 17	.06	.06	0	.06	.06	0
Age $\leq$ 16	.00	.00	0	.00	.00	0

Table 26

$\emptyset$ and $\emptyset/\emptyset_{max}$ for Self-Reported and Computed Age

Group	Coefficient	Age in Years 22 or more	21	20	19	18	17	16 or less
All men	$\emptyset$	.88	.87	.87	.92	.92	.96	--
	$\emptyset/\emptyset_{max}$	.89	.89	.94	.95	.93	.97	--
All women	$\emptyset$	.77	.71	.83	.93	.90	.92	--
	$\emptyset/\emptyset_{max}$	.77	.98	.93	.96	.91	.93	--

Table 27

Correlations Between Item Responses Processed by Optical Scanning and by Verified Keypunching

Item Types	Range of Correlations	Median Correlations
Demographic characteristics	.97--1.00	.99
Financial Concern and Sources	.95-- .96	.96
Career Choice and Major Field	.99	.99
Parents' characteristics	.88-- .98	.93
High school accomplishments	.98--1.00	.99
High school activities	.91--1.00	.95
Goals	.92-- .98	.95
Attitudes	.90-- .98	.93
Expectancies	.90-- .98	.95

Appendix A

1969 Student Information Form

1969 STUDENT INFORMATION FORM

519294

YOUR NAME (please print) _____
First Middle or Maiden Last

HOME STREET ADDRESS _____

City State Zip Code (if known)

When were you born? Month (01-12) Day (01-31) Year

DO NOT MARK THIS GRID

Dear Student:

The information in this report is being collected as part of a continuing study of higher education by the American Council on Education. The Council, which is a non-governmental association of colleges and educational organizations, is soliciting your cooperation in this research in order to achieve a better understanding of how students are affected by their college experiences. Identifying information has been requested in order to make subsequent mail follow-up studies possible. Your response will be held in the strictest professional confidence, and will be used only in group summaries for research purposes.

Sincerely yours,
Logan Wilson
Logan Wilson President

DIRECTIONS: Your responses will be read by an optical mark reader. Your careful observance of these few simple rules will be most appreciated.

Use only black lead pencil (No 2½ or softer). Make heavy black marks that fill the circle. Erase cleanly any answer you wish to change. Make no stray markings of any kind.

	Yes	No
EXAMPLE: Will marks made with ball pen or fountain pen be properly read?	○	●

1. Your Sex: Male ○ ..... Female ○

2. How old will you be on December 31 of this year? (Mark one)
   - 16 or younger ○
   - 17 ○
   - 18 ○
   - 19 ○
   - 20 ○
   - 21 ○
   - Older than 21 ○

3. What was your average grade in secondary school? (Mark one)
   - A or A+ ○
   - A- ○
   - B+ ○
   - B ○
   - B- ○
   - C+ ○
   - C ○
   - D ○

4. To how many colleges other than this one did you actually apply for admission? From how many did you receive acceptances? (Mark one in each column)

	Applications	Acceptances
No other	○	○
One	○	○
Two	○	○
Three	○	○
Four	○	○
Five	○	○
Six or more	○	○

5. Mark one:
   - This is the first time I have enrolled in college as a freshman ○
   - I came to this college from a two-year college ○
   - I came to this college from a four-year college or university ○

6. The following questions deal with accomplishments that might possibly apply to your high school years. Do not be discouraged by this list; it covers many areas of interest and few students will be able to say "yes" to many items. (Mark all that apply)

	Yes
Was elected president of one or more student organizations (recognized by the school)	○
Received a high rating (Good, Excellent) in a state or regional music contest	○
Participated in a state or regional speech or debate contest	○
Had a major part in a play	○
Won a varsity letter (sports)	○
Won a prize or award in an art competition	○
Edited the school paper, yearbook, or literary magazine	○
Had poems, stories, essays, or articles published	○
Participated in a National Science Foundation summer program	○
Placed (first, second, or third) in a state or regional science contest	○
Was a member of a scholastic honor society	○
Won a Certificate of Merit or Letter of Commendation in the National Merit Program	○

7. What is the highest academic degree that you intend to obtain? That your parents hope you will obtain? (Mark one in each column)

	You Intend	Your Parents Hope
None	○	○
Associate (or equivalent)	○	○
Bachelor's degree (B.A., B.S., etc.)	○	○
Master's degree (M.A., M.S., etc.)	○	○
Ph.D or Ed.D.	○	○
M.D., D.D.S., or D.V.M.	○	○
LL.B. or J.D.	○	○
B.D.	○	○
Other	○	○

8. Do you have any concern about your ability to finance your college education? (Mark one)
   - None (I am confident that I will have sufficient funds) ○
   - Some concern (but I will probably have enough funds) ○
   - Major concern (not sure I will be able to complete college) ○

9. Are you a U.S. Citizen? (Mark one)
   - Yes, native born ○
   - Yes, naturalized ○
   - No ○

10. Through what source do you intend to finance the first year of your undergraduate education? (Mark one in each row)

	Major Source	Minor Source	Not a Source
Personal savings and or employment	O	O	O
Parental or other family aid	O	O	O
Repayable loan	O	O	O
Scholarship, grant, or other gift	O	O	O

11. What is the highest level of formal education obtained by your parents? (Mark one in each column)

	Father	Mother
Grammar school or less	O	O
Some high school	O	O
High school graduate	O	O
Some college	O	O
College degree	O	O
Postgraduate degree	O	O

12. What is your best estimate of the total income last year of your parental family (not your own family if you are married)? Consider annual income from all sources before taxes. (Mark one)

Less than $4,000	O
$4,000-$5,999	O
$6,000-$7,999	O
$8,000-$9,999	O
$10,000-$14,999	O
$15,000-$19,999	O
$20,000-$24,999	O
$25,000-$29,999	O
$30 000 or more	O

13. What is your racial background? (Mark one)

White/Caucasian	O
Black/Negro/Afro-American	O
American Indian	O
Oriental	O
Other	O

14. Mark one in each column:

	Religion in Which you Were Reared	Your Present Religious Preference
Baptist	O	O
Congregational (United Church of Christ)	O	O
Episcopal	O	O
Jewish	O	O
Latter Day Saints (Mormon)	O	O
Lutheran	O	O
Methodist	O	O
Muslim	O	O
Presbyterian	O	O
Quaker (Society of Friends)	O	O
Roman Catholic	O	O
Seventh Day Adventist	O	O
Unitarian-Universalist	O	O
Other Protestant	O	O
Other Religions	O	O
None	O	O

15. Where did you rank academically in your high school graduating class? (Mark one)

Top 1%	Top 10%	Top Quarter
2nd Quarter	3rd Quarter	4th Quarter

16. During the next few years, to what extent do you think the Federal Government should be involved in each of the following national issues? (Mark one in each row)

1. Eliminate any existing programs or remain uninvolved
2. Decrease involvement from current levels
3. Maintain current level of involvement
4. Increase involvement from current level
5. Initiate new crash program

1 2 3 4 5	Control of cigarette advertising
1 2 3 4 5	Elimination of violence from T.V.
1 2 3 4 5	Control of environmental pollution
1 2 3 4 5	Use of tax incentives to control the birth rate
1 2 3 4 5	Protection of the consumer from faulty goods and services
1 2 3 4 5	Compensatory education for the disadvantaged
1 2 3 4 5	Special benefits for veterans
1 2 3 4 5	Control of firearms
1 2 3 4 5	Elimination of poverty
1 2 3 4 5	Crime prevention
1 2 3 4 5	School desegregation
1 2 3 4 5	Compensatory financial aid for the disadvantaged
1 2 3 4 5	Control of student activists

17. Where did you live for most of the time while you were growing up?

On a farm	O
In a small town	O
In a moderate size town or city	O
In a suburb of a large city	O
In a large city	O

MAKE SURE YOU HAVE ANSWERED ITEMS 18-21

18. Mark one in each row:

1. Left
2. Liberal
3. Middle-of-the-road
4. Moderately conservative
5. Strongly conservative

1 2 3 4 5	How would you characterize yourself politically at the present time?
1 2 3 4 5	How do you think you will characterize yourself politically four years from now?

19. How many miles is this college from your home?

10 or less	O	101-500	O
11-50	O	501-1000	O
51-100	O	More than 1000	O

20. From what kind of secondary school did you graduate? (Mark one)

Public	O
Private (denominational)	O
Private (non-denominational)	O
Other	O

21. How many brothers and sisters now living do you have? (Mark one in each row)

	None	1	2	3 or more
Number of older brothers	O	O	O	O
Number of older sisters	O	O	O	O
Number of younger brothers	O	O	O	O
Number of younger sisters	O	O	O	O

22. What is your best guess as to the chances that you will: (Mark one in each row)

	Very Good Chance	Some Chance	Very Little Chance	No Chance
Get married while in college?	O	O	O	O
Get married within a year after college?	O	O	O	O
Obtain an A- or better over-all grade point average?	O	O	O	O
Change major field?	O	O	O	O
Change career choice?	O	O	O	O
Fail one or more courses?	O	O	O	O
Graduate with honors?	O	O	O	O
Be elected to a student office?	O	O	O	O
Join a social fraternity, sorority, or club?	O	O	O	O
Author or co-author a published article?	O	O	O	O
Be drafted while I am in college?	O	O	O	O
Be elected to an academic honor society?	O	O	O	O
Protest against U.S. military policy?	O	O	O	O
Protest against administrative policy at this college?	O	O	O	O
Protest against existing racial or ethnic policies?	O	O	O	O
Drop out of this college temporarily (exclude transferring)?	O	O	O	O
Enlist in the Armed Services before graduating?	O	O	O	O
Be more successful after graduation than most students attending this college?	O	O	O	O
Drop out permanently (exclude transferring)?	O	O	O	O
Transfer to another college before graduating?	O	O	O	O

23. Mark one in each column:

	Your current home state	Your birthplace	Father's birthplace	Mother's birthplace
Alabama	○	Y	F	M
Alaska	○	Y	F	M
Arizona	○	Y	F	M
Arkansas	○	Y	F	M
California	○	Y	F	M
Colorado	○	Y	F	M
Connecticut	○	Y	F	M
Delaware	○	Y	F	M
D.C.	○	Y	F	M
Florida	○	Y	F	M
Georgia	○	Y	F	M
Hawaii	○	Y	F	M
Idaho	○	Y	F	M
Illinois	○	Y	F	M
Indiana	○	Y	F	M
Iowa	○	Y	F	M
Kansas	○	Y	F	M
Kentucky	○	Y	F	M
Louisiana	○	Y	F	M
Maine	○	Y	F	M
Maryland	○	Y	F	M
Massachusetts	○	Y	F	M
Michigan	○	Y	F	M
Minnesota	○	Y	F	M
Mississippi	○	Y	F	M
Missouri	○	Y	F	M
Montana	○	Y	F	M
Nebraska	○	Y	F	M
Nevada	○	Y	F	M
New Hampshire	○	Y	F	M
New Jersey	○	Y	F	M
New Mexico	○	Y	F	M
New York	○	Y	F	M
North Carolina	○	Y	F	M
North Dakota	○	Y	F	M
Ohio	○	Y	F	M
Oklahoma	○	Y	F	M
Oregon	○	Y	F	M
Pennsylvania	○	Y	F	M
Rhode Island	○	Y	F	M
South Carolina	○	Y	F	M
South Dakota	○	Y	F	M
Tennessee	○	Y	F	M
Texas	○	Y	F	M
Utah	○	Y	F	M
Vermont	○	Y	F	M
Virginia	○	Y	F	M
Washington	○	Y	F	M
West Virginia	○	Y	F	M
Wisconsin	○	Y	F	M
Wyoming	○	Y	F	M
Canada	○	Y	F	M
Mexico	○	Y	F	M
Puerto Rico	○	Y	F	M
Other Latin America	○	Y	F	M
Europe	○	Y	F	M
Asia	○	Y	F	M
Other	○	Y	F	M

24. Mark only three responses, one in each column.

Your probable career occupation.
Your father's occupation.
Your mother's occupation.

Y F M

NOTE: If your father (or mother) is deceased, please indicate his (her) last occupation.

	Y	F	M
Accountant or actuary	Y	F	M
Actor or entertainer	Y	F	M
Architect	Y	F	M
Artist	Y	F	M
Business (clerical)	Y	F	M
Business executive (management, administrator)	Y	F	M
Business owner or proprietor	Y	F	M
Business salesman or buyer	Y	F	M
Clergyman (minister, priest)	Y	F	M
Clergy (other religious)	Y	F	M
Clinical psychologist	Y	F	M
College teacher	Y	F	M
Computer programmer	Y	F	M
Conservationist or forester	Y	F	M
Dentist (including orthodontist)	Y	F	M
Dietitian or home economist	Y	F	M
Engineer	Y	F	M
Farmer or rancher	Y	F	M
Foreign service worker (including diplomat)	Y	F	M
Housewife	Y	F	M
Interior decorator (including designer)	Y	F	M
Interpreter (translator)	Y	F	M
Lab technician or hygienist	Y	F	M
Law enforcement officer	Y	F	M
Lawyer (attorney)	Y	F	M
Military service (career)	Y	F	M
Musician (performer, composer)	Y	F	M
Nurse	Y	F	M
Optometrist	Y	F	M
Pharmacist	Y	F	M
Physician	Y	F	M
School counselor	Y	F	M
School principal or superintendent	Y	F	M
Scientific researcher	Y	F	M
Social worker	Y	F	M
Statistician	Y	F	M
Therapist (physical, occupational, speech)	Y	F	M
Teacher (elementary)	Y	F	M
Teacher (secondary)	Y	F	M
Veterinarian	Y	F	M
Writer or journalist	Y	F	M
Skilled trades	Y	F	M
Other	Y		
Undecided	Y		
Laborer (unskilled)		F	M
Semi-skilled worker		F	M
Other occupation		F	M
Unemployed		F	

25. In general, my parents could be characterized as:
(Mark one circle in each row.)

1. Strongly agree
2. Agree
3. Disagree
4. Strongly disagree

	1	2	3	4
Interested in intellectual pursuits	1	2	3	4
Interested in cultural pursuits	1	2	3	4
Deeply religious	1	2	3	4
Interested in politics	1	2	3	4
Deeply concerned about their children	1	2	3	4
Financially comfortable	1	2	3	4
Having high aspirations for me	1	2	3	4

26. Below is a general list of things that students sometimes do. Indicate which of these things you did during the past year in school. If you engaged in an activity frequently, mark "F". If you engaged in an activity one or more times, but not frequently, mark "O" (occasionally). Mark "N" (not at all) if you have not performed the activity during the past year. (Mark one for each item)

	Frequently	Occasionally	Not at all
Voted in a student election	F	O	N
Came late to class	F	O	N
Played a musical instrument	F	O	N
Studied in the library	F	O	N
Checked out a book or journal from the school library	F	O	N
Arranged a date for another student	F	O	N
Overslept and missed a class or appointment	F	O	N
Typed a homework assignment	F	O	N
Discussed my future with my parents	F	O	N
Failed to complete a homework assignment on time	F	O	N
Argued with a teacher in class	F	O	N
Attended a religious service	F	O	N
Demonstrated for a change in some racial or ethnic policy	F	O	N
Demonstrated for a change in some military policy	F	O	N
Demonstrated for a change in some administrative policy of my high school	F	O	N
Did extra (unassigned) reading for a course	F	O	N
Took sleeping pills	F	O	N
Tutored another student	F	O	N
Played chess	F	O	N
Read poetry not connected with a course	F	O	N
Took a tranquilizing pill	F	O	N
Discussed religion	F	O	N
Took vitamins	F	O	N
Visited an art gallery or museum	F	O	N
Worked in a school political campaign	F	O	N
Worked in a local, state, or national political campaign	F	O	N
Missed school because of illness	F	O	N
Smoked cigarettes	F	O	N
Discussed politics	F	O	N
Drank beer	F	O	N
Discussed sports	F	O	N
Asked a teacher for advice after class	F	O	N
Had vocational counseling	F	O	N
Stayed up all night	F	O	N

27. Below is a list of 66 different undergraduate major fields grouped into general categories. Mark only three of the 66 fields as follows:

① First choice (your probable major field of study).
② Second choice.
Ⓛ The field of study which is least appealing to you.

ARTS AND HUMANITIES

Architecture	① ② Ⓛ
English (literature)	① ② Ⓛ
Fine arts	① ② Ⓛ
History	① ② Ⓛ
Journalism (writing)	① ② Ⓛ
Language (modern)	① ② Ⓛ
Language (other)	① ② Ⓛ
Music	① ② Ⓛ
Philosophy	① ② Ⓛ
Speech and drama	① ② Ⓛ
Theology	① ② Ⓛ
Other	① ② Ⓛ

BIOLOGICAL SCIENCE

Biology (general)	① ② Ⓛ
Biochemistry	① ② Ⓛ
Biophysics	① ② Ⓛ
Botany	① ② Ⓛ
Zoology	① ② Ⓛ
Other	① ② Ⓛ

BUSINESS

Accounting	① ② Ⓛ
Business admin	① ② Ⓛ
Electronic data processing	① ② Ⓛ
Secretarial studies	① ② Ⓛ
Other	① ② Ⓛ

ENGINEERING

Aeronautical	① ② Ⓛ
Civil	① ② Ⓛ
Chemical	① ② Ⓛ
Electrical	① ② Ⓛ
Industrial	① ② Ⓛ
Mechanical	① ② Ⓛ
Other	① ② Ⓛ

PHYSICAL SCIENCE

Chemistry	① ② Ⓛ
Earth science	① ② Ⓛ
Mathematics	① ② Ⓛ
Physics	① ② Ⓛ
Statistics	① ② Ⓛ
Other	① ② Ⓛ

PROFESSIONAL

Health Technology (medical, dental, laboratory)	① ② Ⓛ
Nursing	① ② Ⓛ
Pharmacy	① ② Ⓛ
Predentistry	① ② Ⓛ
Prelaw	① ② Ⓛ
Premedical	① ② Ⓛ
Preveterinary	① ② Ⓛ
Therapy (occupat., physical, speech)	① ② Ⓛ
Other	① ② Ⓛ

SOCIAL SCIENCE

Anthropology	① ② Ⓛ
Economics	① ② Ⓛ
Education	① ② Ⓛ
History	① ② Ⓛ
Political science (government, int. relations)	① ② Ⓛ
Psychology	① ② Ⓛ
Social work	① ② Ⓛ
Sociology	① ② Ⓛ
Other	① ② Ⓛ

OTHER FIELDS

Agriculture	① ② Ⓛ
Communications (radio, T.V., etc.)	① ② Ⓛ
Electronics (technology)	① ② Ⓛ
Forestry	① ② Ⓛ
Home economics	① ② Ⓛ
Industrial arts	① ② Ⓛ
Library science	① ② Ⓛ
Military science	① ② Ⓛ
Physical education and recreation	① ② Ⓛ
Other (technical)	① ② Ⓛ
Other (nontechnical)	① ② Ⓛ
Undecided	① ② Ⓛ

Please be sure that only three circles have been marked in the above list.

28. Indicate the importance to you personally of each of the following: (Mark one for each item)

	Essential	Very important	Somewhat important	Not important
Becoming accomplished in one of the performing arts (acting, dancing, etc.)	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Becoming an authority on a special subject in my subject field	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Obtaining recognition from my colleagues for contributions in my special field	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Influencing the political structure	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Influencing social values	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Raising a family	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Having an active social life	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Having friends with different backgrounds and interests from mine	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Becoming an expert in finance and commerce	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Having administrative responsibility for the work of others	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Being very well-off financially	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Helping others who are in difficulty	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Becoming a community leader	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Making a theoretical contribution to science	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Writing original works (poems, novels, short stories, etc.)	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Never being obligated to people	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Creating artistic work (painting, sculpture, decorating, etc.)	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Keeping up to date with political affairs	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Being successful in a business of my own	Ⓔ	Ⓥ	Ⓢ	Ⓝ
Developing a meaningful philosophy of life	Ⓔ	Ⓥ	Ⓢ	Ⓝ

29. Mark one in each row:
Agree strongly
Agree somewhat
Disagree somewhat
Disagree strongly

	Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
Students should have a major role in specifying the college curriculum	O	O	O	O
Scientists should publish their findings regardless of the possible consequences	O	O	O	O
Realistically, an individual person can do little to bring about changes in our society	O	O	O	O
College officials have the right to regulate student behavior off campus	O	O	O	O
The chief benefit of a college education is that it increases one's earning power	O	O	O	O
Faculty promotions should be based in part on student evaluations	O	O	O	O
My beliefs and attitudes are similar to those of most other students	O	O	O	O
Student publications should be cleared by college officials	O	O	O	O
Marijuana should be legalized	O	O	O	O
College officials have the right to ban persons with extreme views from speaking on campus	O	O	O	O
Only volunteers should serve in the armed forces	O	O	O	O
Students from disadvantaged social backgrounds should be given preferential treatment in college admissions	O	O	O	O
Most college officials have been too lax in dealing with student protests on campus	O	O	O	O
Divorce laws should be liberalized	O	O	O	O
Under some conditions, abortions should be legalized	O	O	O	O
There is too much concern in the courts for the rights of criminals	O	O	O	O
Capital punishment (the death penalty) should be abolished	O	O	O	O

Prepared by Office of Research, American Council on Education 1785 Massachusetts Ave., N.W. Washington, D.C. Processed By National Computer Systems 1015 South 6th Street Mpls., Minn. 55415

Appendix B

Supplemental Student Information Form

SUPPLEMENTAL STUDENT INFORMATION FORM

NAME _____ COLLEGE _____

TESTING # _____

DIRECTIONS: This form is a continuation of the form you just completed.

Please continue to indicate your answers by making heavy black marks that fill the circle. Continue to use only the black lead pencil.

When you have completed both forms, make sure that your name and testing number is on both forms. Return the completed forms to the test administrator.

1. How would you rate the academic standards of your high school? (Mark one)

Very high .....O
Fairly high .....O
About average.....O
Probably below average .....O
Definitely below average.....O

2. To what extent do you think each of the following describes the psychological climate or atmosphere at this college? (Mark one answer for each item)

	Very Descriptive	In Between	Not at all Descriptive
Intellectual	O	O	O
Snobbish	O	O	O
Social	O	O	O
Victorian	O	O	O
Practical-minded	O	O	O
Warm	O	O	O
Realistic	O	O	O
Liberal	O	O	O

3. Answer each of the following as you think it applies to this college:

	Yes	No
The students are under a great deal of pressure to get high grades	O	O
The student body is apathetic and has little "school spirit"	O	O
Most of the students are of a very high calibre academically	O	O
There is a keen competition among most of the students for high grades	O	O
Freshmen have to take orders from upperclassmen for a period of time	O	O
There isn't much to do except to go to class and study	O	O
I felt "lost" when I first came to the campus	O	O
Being in this college builds poise and maturity	O	O
Athletics are overemphasized	O	O
The classes are usually run in a very informal manner	O	O
Most students are more like "numbers in a book"	O	O

4. Indicate the importance to you personally of the following persons or events in your decision to enroll in this college. (Mark one for each item)

	Major Influence	Minor Influence	Not Relevant
Parent or other relative	O	O	O
High school teacher or counselor	O	O	O
Friends attending this college	O	O	O
Graduate or other representative from this college	O	O	O
Professional counseling or college placement service	O	O	O
Athletic program of the college	O	O	O
Other extracurricular activities	O	O	O
Social life of the college	O	O	O
Opportunity to live away from home	O	O	O
Low cost	O	O	O
Academic reputation of the college	O	O	O
Most of the students are like me	O	O	O
Religious affiliation	O	O	O

5. Below is a general list of things that students sometime do. Indicate which of these things you did during your senior year in high school. (Mark one for each item)

	Frequently	Occasionally	Not at all
Was a guest in a teacher's home	O	O	O
Rode on a motorcycle	O	O	O
Studied with other students	O	O	O
Saw a foreign movie	O	O	O
Took a trip of more than 500 miles	O	O	O
Got a traffic ticket	O	O	O
Played tennis	O	O	O
Played bridge	O	O	O
Gambled with cards or dice	O	O	O
Drove a car	O	O	O
Attended a ballet performance	O	O	O
Participated on the speech or debate team	O	O	O
Acted in plays	O	O	O
Sang in a choir or glee club	O	O	O
Called a teacher by his or her first name	O	O	O
Wrote an article for the school paper or literary magazine	O	O	O
Had a blind date	O	O	O
Wrote a short story or poem (not for a class)	O	O	O
Discussed how to make money with other students	O	O	O
Attended a public recital or concert	O	O	O
Went to an over-night or week-end party	O	O	O
Took weight-reducing or dietary formula	O	O	O
Participated in an informal group sing	O	O	O
Drank wine	O	O	O
Cribbed on an examination	O	O	O
Turned in a paper or theme late	O	O	O
Asked questions in class	O	O	O

6. Mark one in each row:

Agree strongly
Agree somewhat
Disagree somewhat
Disagree strongly

	Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
Current levels of air pollution in large cities justify the use of drastic measures to limit the use of motor vehicles	O	O	O	O
Urban problems cannot be solved without huge investments of Federal money	O	O	O	O
Cigarette advertising should be outlawed on radio and T.V.	O	O	O	O
College faculty are more competent than are students to specify the curriculum	O	O	O	O
The activities of married women are best confined to the home and family	O	O	O	O
Parents should be discouraged from having large families	O	O	O	O
Colleges would be improved if organized sports were de-emphasized	O	O	O	O
Women should be subject to the draft	O	O	O	O
The voting age should be lowered to 18	O	O	O	O

7. Rate yourself on each of the following traits as you really think you are when compared with the average student of your own age. We want the most accurate estimate of how you see yourself. (Mark one for each item)

Trait	Highest 10 Percent	Above Average	Average	Below Average	Lowest 10 Percent
Academic ability	O	O	O	O	O
Athletic ability	O	O	O	O	O
Artistic ability	O	O	O	O	O
Cheerfulness	O	O	O	O	O
Defensiveness	O	O	O	O	O
Drive to achieve	O	O	O	O	O
Leadership ability	O	O	O	O	O
Mathematical ability	O	O	O	O	O
Mechanical ability	O	O	O	O	O
Originality	O	O	O	O	O
Political conservatism	O	O	O	O	O
Political liberalism	O	O	O	O	O
Popularity	O	O	O	O	O
Popularity with the opposite sex	O	O	O	O	O
Public speaking ability	O	O	O	O	O
Self-confidence (intellectual)	O	O	O	O	O
Self-confidence (social)	O	O	O	O	O
Sensitivity to criticism	O	O	O	O	O
Stubbornness	O	O	O	O	O
Understanding of others	O	O	O	O	O
Writing ability	O	O	O	O	O

Other Recent Publications by the Staff of the Office of Research

American Council on Education
(ACE)

Astin, A. W. **College Dropouts: A National Profile.** ACE Research Reports, Vol. 7, No. 1, Washington: ACE, 1972.

Astin, A. W. *New Evidence on Campus Unrest, 1969-70.* **Educational Record, 52** (Winter 1971), 41-46.

Astin, A. W. *Evaluative Research and the New Colleges.* In P. Dressel (Ed.), **The New Colleges: Toward an Appraisal.** Iowa City: American College Testing Program and the American Association for Higher Education, 1971.

Astin, A. W. **Predicting Academic Performance In College.** New York: Free Press, 1971.

Astin, A. W., and Lee, C. B. T. **The Invisible Colleges.** Carnegie Commission Series on Higher Education. New York: McGraw-Hill Book Co., 1971.

Astin, H. S., and Bayer, A. E. *Sex Discrimination in Academe.* **Educational Record, 53** (Spring 1972), 101-18.

Bayer, A. E. **College and University Faculty: A Statistical Description.** ACE Research Reports, Vol. 5, No. 5. Washington: ACE, 1971.

Bayer, A. E. *Institutional Correlates of Faculty Support of Campus Unrest.* **Sociology of Education, 45** (Winter 1972), 76-94.

Bayer, A. E. *Faculty Roles in Campus Unrest.* **Change Magazine, 3** (Winter 1971-72), 10, 74-75.

Creager, J. A. *Academic Achievement and Institutional Environments: Two Research Strategies.* **Journal of Experimental Education, 40** (Winter 1971), 9-23.

Creager, J. A. *Futurism in Higher Education.* **Change Magazine, 4** (Winter 1972), 8, 62.

Creager, J. A. *Orthogonal and Nonorthogonal Methods for Partitioning Regression Variance.* **American Educational Research Journal, 8** (November 1971), 671-76.

Creager, J. A. **The American Graduate Student: A Normative Description.** ACE Research Reports, Vol. 6, No. 5. Washington: ACE, 1971.

Creager, J. A. **Goals and Achievements of the ACE Academic Internship Program.** ACE Research Reports, Vol. 6, No. 3, Washington: ACE, 1971.

Drew, D. E. *Jewish Students Today: Radical or Conservative?"* **Transaction, 8** (October 1971), 45-48.

Drew, D. E. **A Study of the NSF College Science Improvement Program.** ACE Research Reports, Vol. 6, No. 4. Washington: ACE, 1971.

Greeley, A. M., Astin, A. W., Katz, J., and Epperson, D. C. *The Newman Report: Four Comments.* **Journal of Higher Education, 42** (October 1971), 610-23.

Higher Education Panel Survey. *Freshman Class Vacancies in Fall 1971 and Recent Trends in Enrollments of Minority Freshmen.* Spring 1972. Mimeographed.

Higher Education Panel Survey. *What College Presidents Are Reading.* Spring 1972. Mimeographed.

Karabel, J. B. *Perspectives on Open Admissions.* **Educational Record, 53** (Winter 1972), 30-44.

Rossmann, J. E. *Interest Patterns Among Economists.* **Journal of Counseling Psychology, 19** (May 1971), 255-61.

Rossmann, J. E. *The Interim Term After Seven Years.* **Journal of Higher Education, 42** (October 1971) 603-09.

Rossmann, J. E., Johansson, C., and Sandell, S. *The S-U Grading Option.* **Educational Record, 52** (Summer 1971), 273-76.

Staff of the Office of Research. **The American Freshman: National Norms for Fall 1971.** ACE Research Reports, Vol. 6, No. 6. Washington: ACE, 1971.