DOCUMENT RESUME

ED 096 193 95 SO 007 712

AUTHOR Knudsen, Kjell R.; And Others
TITLE Emphasis in Educational Research: A Description of Selected Substantive and Methodological Characteristics of Proposals in the 1973 National Institute of Education (NIE) Field Initiated Studies (FIS) Program.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 16 Apr 74
CONTRACT NE-C-00-3-0287
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, April 1974); The project on which this paper is based was conducted by Minnesota Systems Research, Inc., Minneapolis

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Content Analysis; *Educational Research; Evaluation Methods; Federal Aid; *Federal Programs; *Research Methodology; *Research Proposals; Research Utilization
IDENTIFIERS Field Initiated Studies; *National Institute of Education; NIE

ABSTRACT

Some 3,000 proposals submitted for Research Grants in Education and Small Grants, two of the major topic areas in the 1973 Field Initiated Studies (FIS) program of the National Institute of Education (NIE), are analyzed in this report to identify research features which seemed to receive priority emphasis and to describe some of the methodological features of these projects in order to focus questions about the potential for integrated, cumulative knowledge in education. The report contains selected characteristics of research from a considerably longer list in the original report which is the third in a related series of investigations in which the characteristics of federally funded social research are assessed. Objectives of the study for NIE were to develop an instrument to assess the methodological quality of the research component of proposals; describe the substantive issues of proposals; determine issues removed and retained at each step of the NIE review process; relate proposed characteristics of the research component; and assess systematic preferences in the NIE proposal review process in terms of the review process used by Minnesota Systems Research, which conducted the project. (Author/JH)

EMPHASIS IN EDUCATIONAL RESEARCH: A DESCRIPTION OF SELECTED SUBSTANTIVE AND METHODOLOGICAL CHARACTERISTICS OF PROPOSALS IN THE 1973 NATIONAL INSTITUTE OF EDUCATION (NIE) FIELD INITIATED STUDIES (FIS) PROGRAM

Presented at the 1974 Annual Meeting of the American Educational Research Association April 16, 1974, Chicago, Illinois

KJELL R KNUDSEN, PhD
Minnesota Systems Research, Inc

DONALD G McTAVISH, PhD
University of Minnesota

EDWARD E BRENT, Jr
Minnesota Systems Research, Inc

JAMES D CLEARY
University of Minnesota

The study from which these data come was supported by contract number NE-C-00-3-0287 from the National Institute of Education. Dr. Willy De Geyndt was the principal investigator and Dr Kjell R Knudsen was the co-principal investigator and project director. Mr Edward E. Brent and Mr. James D. Cleary are both doctoral candidates in Sociology at the University of Minnesota and Dr Donald McTavish is on the faculty in Sociology at the University of Minnesota, Minneapolis. The project was conducted through Minnesota Systems Research, Inc, a non-profit, research organization, Minneapolis, Minnesota. The final report of this study, including technical notes on methodology and procedures, was submitted in February, 1974 to the National Institute of Education.

The conclusions and recommendations expressed in this paper should not be interpreted as expressing or representing the opinion or policy of the National Institute of Education or Minnesota Systems Research, Inc.

INTRODUCTION

Potentially, educational research involves an unusually broad range of theoretical perspectives, units of analysis and types of variables due in part to the number of specific disciplines which are relevant. This variety poses critical problems for interrelating research results and providing for the cumulative development of knowledge in education. Analytic discussions of the state of the field are helpful in this regard,[1] pointing out overlaps, contradictions and unanswered questions in past research.[2] But what increment in knowledge might we expect from the current investment in research which will be completed in the near future? In what directions are we likely to go? With what kind of emphasis and balance? What variables and topics are being examined and are these indeed the areas most in need of research? To what extent does this research have methodological features which will maximize its potential for making a cumulative addition to systematic knowledge in education? Little empirical description of these trends and potentials is available in current literature.

The following discussion provides information about characteristics of a relatively large collection of recently proposed and funded educational research, the 1973 NIE Field Initiated Studies (FIS) program. While we make no claim that this collection of research is representative of all current educational research, we do feel that this program has certain features which may help provide insight into some of these questions. It was a program which distributed some $10 million in research support, an amount which is substantial enough to have a large potential impact. Some three thousand proposals were submitted, largely on topics considered relevant by those attuned to this funding source. Professionals both within government and outside made selection judgments and funded nearly two-hundred projects. These were professionals generally in a position to be aware of educational research trends and needs. Our description, then, is intended to help identify some of the features of this research which seemed to receive priority emphasis, and to describe some of the methodological features of these projects in order to focus questions about the potential for integrated, cumulative knowledge in education. This paper presents only selected characteristics

of research from a considerably longer list in our original report. This study is the third of a related series of investigations in which we have been involved to assess characteristics of federally funded social research.[3] Previous studies in this series have focused primarily on methodological features of ongoing and completed projects. In this study, only research proposals were examined. This means that some features of research could only be judged on whether they were 'planned' or 'likely' rather than actually accomplished.

The particular objectives for the study for NIE were to: (a) develop an instrument to assess the methodological quality of the research component of proposals; (b) describe the substantive issues of proposals in appropriate categories; (c) determine issues removed and retained at each step of the NIE review process; (d) relate proposer characteristics to preferences for substantive issues and characteristics of the research component of proposals; and (e) assess systematic preferences in the NIE proposal review process in terms of the review process used by Minnesota Systems Research, Inc (MSRI).

Our focus in this paper is on the 2,951 proposals submitted for Research Grants in Education and Small Grants, two of the major topic areas in the 1973 FIS program. Excluded from this paper are proposals in the Selected Discipline areas, which were reviewed in somewhat different steps. The steps through which a successful proposal travelled are shown in Figure 1:

FIGURE 1

STEPS IN FIS 1973 FOR RESEARCH GRANTS IN EDUCATION AND SMALL GRANTS

1 Request for proposals mailed to approximately 25,000 people
2 Screening of proposals by in-house personnel
3 Panel review of proposals by outside professionals
4 Funding decision based on panel recommendations

As can be seen from the figure, the NIE review process consisted of screeners and panels. Practically all proposals approved by the panels were funded by the Institute.

Blocks of variables examined in the course of the study were:

1. Descriptive Methodology
2. Characteristics of Principal Investigators
3. Descriptive Content
4. Evaluative Methodology

These were measured by our two instruments and these blocks were the basis of models analyzed in the larger study. Here, we will present selected outcomes in terms of these blocks of variables.

METHODOLOGY

The evaluation study design called for a set of research professionals to rate a stratified, non-proportional, random sample of proposals from the population. The population and sample are shown in Table 1. Two instruments were constructed for use by the raters to rate the proposals. One instrument dealt with a conceptual analysis of research proposals, including proposer characteristics and descriptive content variables. The other instrument dealt with research components and was oriented toward descriptive and evaluative methodological variables. A substantial amount of information was, thus, available (over 500 variables) in the data base.

Proposals were randomly assigned to five raters who had been hired as summer interns by NIE. All raters had relatively strong backgrounds in research methods and experience in doing research. Three of the raters were in sociology and two in human development and all had some professional interest in the field of education. All were either in doctoral programs or had relatively recently received their PhD degree. The raters received training before rating proposals and they were instructed to work independently on the rating task. They were located at NIE in Washington, DC. A rater

TABLE 1 - NUMBER OF PROJECTS IN THE POPULATION AND STUDY SAMPLE BY DECISION STEP IN PROCESSING PROPOSED GRANTS FOR RESEARCH IN EDUCATION AND SMALL GRANTS SUBMITTED FOR NIE FIS 1973

Number of Projects in:	Rejected by Screeners	Rejected by Panel	Funded	Total Proposals
FIS 1973 Population	2345	458	148	2951
Study Sample	220	73	148	441

stability and consistency check was designed along with the main rating task.[4] Raters were debriefed at the end of the rating task as a further means of assessing rating norms and definitions. They were not privy to design details of our study until after all rating had been completed.

CHARACTERISTICS OF FUNDED PROPOSALS

Table 2 presents the distribution of selected characteristics of research proposals as measured by our instruments. These are broken down into four general categories corresponding to the blocks of variables shown in the model above. The left-most column in Table 2 is the distribution among all submitted proposals based on the weighted sample. The right-most column is the distribution among proposals ultimately funded. The center two columns express the net effect (difference between those reviewed and those approved) of in-house staff screeners and of the outside, professional panelists.

Over half (54%) of the funded research (and 57% of submitted proposals) was proposed by principal investigators who were in the field of education. Psychologists were second most frequent (36%) and about double their percentage representation among submitted proposals. Approximately five percent of the principal investigators were from minority groups as best this could be judged from limited available information. In terms of substantive topics, the area of development was most frequent among funded projects (55%), with research on cognitive processes (45%) and achievement (45%) close behind. Ability/aptitude, development and cognitive processes were substantive topics which increased their percentage appearance by at least ten percent in the funded compared to the submitted proposals.

Among funded projects, about half investigated some identifiable educational system (53%), a considerable drop from the emphasis found among all submitted proposals (75%). For projects which had an identifiable educational system, we sought to characterize that system in a number of ways. One way involved a series of pairs of opposed characteristics, such as 'established' versus 'new'. Among proposals where an unambiguous decision between the pairs

TABLE 2 - THE DISTRIBUTION OF SELECTED CHARACTERISTICS OF RESEARCH PROPOSALS BEFORE AND AFTER SCREENING AND PANEL DECISIONS, NIE'S FIS PROGRAM 1973*

	Original Population of Proposals Submitted	Difference Between Distribution Among Proposals Seen By and Accepted By:** In-House Screeners	Outside Panels	Funded Proposals
	%			%
A. PROPOSER CHARACTERISTICS				
Discipline of Principal Investigator	56.6%	- 6.1	+ 3.6	54.1%
Education	18.5	+15.3	+ 2.7	36.5
Psychology	6.1	+ 2.3	+ .4	8.8
Sociology	5.3	- .8	+ .2	4.7
Minority Investigator				
B. DESCRIPTIVE CONTENT VARIABLES				
Substantive Topics	42.8	-15.7	+ 7.4	34.5
Instruction	39.4	- 2.4	+ 8.3	45.3
Achievement	30.4	+ 6.9	+ 5.3	42.6
Ability and Aptitude	29.3	+ 7.8	+17.6	54.7
Development	26.7	+13.4	+ 5.2	45.3
Cognitive Processes	24.4	- 7.7	+15.7	32.4
Attitudes and Values	21.1	+ .8	+ 4.5	26.4
Classroom Role Structure	12.3	- 5.0	+ 1.5	8.8
Preparation for Work	10.8	+ 1.2	- .5	11.5
Equality of Educational Opportunity	5.9	+ 1.7	+ 2.5	10.1
Social Problems and Education				
Percentage of Projects Proposing/Researching Educational System of Some Type	74.8	- 8.0	-13.4	53.4
Paired Characteristics of Educational System Proposed/ Researched***				
Educational Goals Instrumental (rather than Expressive)	87.7	- 6.4	+ 7.9	89.2
Education as a Transmitter of Existing Values & Skills (rather than an Agent of Social Change)	81.6	+ 4.3	+ 7.0	92.9
Academic (rather than Vocational)	77.6	+ 4.5	+14.2	96.3
Socialization for Conformity (rather than Autonomy)	58.1	- 5.3	+18.2	71.0
Established (rather than New)	35.0	- 2.0	+33.6	66.6
School Philosophy Typology****				
Free Schools	69.5	+ 8.7	- 3.2	75.0
Open Schools	15.8	- 8.6	+ 8.3	15.6
Traditional	14.7	- .1	- 5.2	9.4
Variables Used in the Research	32.6	- 7.5	+ 1.3	26.4
Effectiveness	25.8	+ 3.1	+13.0	41.9
Behaviors	15.8	- .6	+13.2	28.4
Race-Ethnicity	15.0	+ 5.3	+ 3.3	23.6
Growth and Development	4.2	+ 1.4	- 3.6	2.0
Creativity				
C. DESCRIPTIVE METHODOLOGICAL VARIABLES				
Classification of Research	41.9	+ 7.1	+ 7.8	56.8
Cross Sectional Study	27.8	+ 9.0	+ 3.7	40.5
Sample Survey	24.4	+22.5	+13.9	60.8
Complex Analysis Techniques	23.5	+ 3.1	+ 5.8	32.4
Qualitative Research	16.7	- .4	+ 4.0	20.3
Case Study	10.0	+ 8.6	+11.1	29.7
Longitudinal Study				
Data Source	56.5	+ 5.5	+ 1.5	63.5
New Questionnaire or Test Data	26.5	+ 9.1	+ .2	35.8
New Observational Data				
D. EVALUATIVE METHODOLOGICAL VARIABLES				
Major Strengths	46.8	- 3.6	+ 2.6	45.9
Applied	21.3	+17.2	+12.9	51.4
Methods	18.5	+13.4	+ 9.3	41.2
Theory	47.7	+12.5	- 4.1	56.1
Judged Likely to Meet Objectives	90.7	- 3.6	+ 6.9	94.0
Data Judged to be Accessible	74.1	+15.8	+ 2.7	92.6
Systematic Analysis Planned				
Methods Scales [Mean, (Standard Deviation)]				
Design (Range 0-5)	1.87 (1.17)	+.70 (.00)	+.29 (-.21)	2.86 (.97)
Measurement Validity (Range 0-5)	.72 (1.13)	+.44 (+.26)	+.21 (-.07)	1.37 (1.32)
Sampling (Range 0-5)	1.37 (1.12)	+.28 (+.16)	+.12 (-.07)	1.77 (1.22)
Statistical Procedures (Range 0-1)	.19 (.24)	+.13 (+.04)	+.09 (.00)	.40 (.29)
Care in Analysis (Range 0-1)	.14 (.25)	+.06 (.00)	+.02 (+.02)	.21 (.27)

*The population, N=2951, includes all proposals submitted to Grants for Research in Education and Small Grants. All of these were seen by in-house screeners. The panels reviewed 606 of the proposals seen by the screeners. NIE funded 148 of the 606 seen by panels. Proposals submitted in the Special Disciplines area (N=143) were omitted, because the review process involved somewhat different steps. Population estimates are based on weighted sample proportions from a sample of 441 proposals.

**A (+) means an increase in percentage from the previous step in the review process; a (-) means a decrease.

***The base of percentages is the number of projects about some identifiable instructional system where the left or right member of each pair could be unambiguously checked. Eliminated are all cases where both could be checked or where the rater was unable to determine which should be checked.

****The base of the percentages is the number of projects which could be classified as belonging in the three types.

was possible, Table 2 provides the percentage emphasizing one side of the pair over the other side, giving some notion of relative emphasis among researched educational systems. Funded research of this type appears to be relatively homogeneous in emphasizing 'established' (67%) rather than 'new', where goals appear to be 'instrumental' (89%) rather than 'expressive', a transmitter of 'existing values and skills' (93%) rather than an agent of 'social change', 'academic' (96%) rather than 'vocational', and aligned toward socialization for 'conformity' (71%) rather than 'autonomy'. More than a ten percent increase in these percentages over that among submitted proposals occurred in all of the above pairs except 'instrumental/expressive'. The educational research community appears to be relatively agreed upon the type of system needing research attention.

One can look at these pairs in a somewhat different way in order to shed some initial light upon areas of emphasis. Two pairs asked for a classification of the researched educational system (if there was one) in terms of whether the teacher's and student's role was 'active' or 'passive'. From this a three-fold typology exhausted all but eight of the potentially classifiable cases. Free school (both active) was by far the most popular focus (75%) among funded projects and submitted projects (70%) with open schools (student not teacher active) and traditional schools (teacher not student active) lagging far behind. This overall emphasis may be due in part to rater's definition of an 'active' role.

In terms of variables used in research, Table 2 indicates that creativity received very little attention (two percent of funded proposals), behaviors received attention in 42% of the funded proposals and effectiveness, race/ethnicity and growth/development received attention in about a quarter of the proposals each. Behaviors and race/ethnicity were considerably more frequent among funded than among all submitted proposals.

Proposed research was more frequently cross-sectional than longitudinal (57% versus 30%), sample survey rather than case study (40% versus 20%), and involved complex analysis techniques (generally more elaborate than

univariate statistics and simple cross tabulations) rather than qualitative research (61% versus 32%) among funded research. Only qualitative and case-study research did not increase substantially in percentage between all proposals and funded proposals. Although proposals involving observational data were substantially more frequent among funded projects than submitted projects, it is clear that questionnaire and test data are still favorites (64% among funded projects)

Finally, in terms of evaluative methodology, 40 to 50 percent of funded proposals were rated to have major strengths in methods, in theory, and in applied areas, and, except for applied strengths, each of these percentages was up substantially (ten percent or more) from the pool of all submitted proposals. Among funded projects, slightly less than 60% were judged to be likely to meet stated goals, nearly all (93%) planned some systematic analysis, and for nearly all (94%) data were judged to be accessible for the study. Systematic analysis seemed preferred by screeners and panelists and as a result, this type of project was more frequent among funded than submitted proposals. In addition to the above ratings, proposals were judged on a relatively large number of specific methodological features which were summarized in terms of five scales: Design, Sampling, Measurement Validity, Statistical Procedures and Care in Analysis (see appended note on scaling for a brief explanation of each scale). All except the latter two range from a low of 0 to a high of 5 (the latter two range from 0 to 1 because only absence or presence of the characteristic was judged as only pronosals were available). On the average, scale scores were higher among funded than submitted proposals. Apparently screeners and panelists were sensitive to educational research needs for improved conceptual clarity as noted by Travers[5] because the design scale showed the greatest average increase and its standard deviation decreased among funded projects as compared to the other scales. In separate analysis we have found that the relative frequency of methodological problems or flaws rises markedly to substantial levels at design scores below about 3.0. The average design score of 2.9 among funded projects is about at this turning point, a substantively important average shift up from the average design score (1.9) among all submitted projects.

Even though a project has an adequate standing on design, there are several other questions about projects which should be addressed, such as: When these projects are completed, are they likely to have addressed key topics? Is their methodology likely to support cumulative knowledge? Is there sufficient variety and are pioneering areas likely to have been supported? If not, why and what steps might one take now to assure a more desired crop of findings in the next five or ten years?

SELECTION EMPHASIS OF IN-HOUSE SCREENERS AND OUTSIDE PANELISTS

The last section characterized educational research in terms of the apparent interests and emphasis among those proposing research and among those select few who would be supported in producing findings to appear in the next two or three years. Another way of assessing the direction of the field is to examine decisions of those who are likely to be most aware of educational research possibilities and needs, both inside a funding agency and experienced professionals from the field. If these two selection groups differ, one might expect that the in-house screeners would emphasize agency goals and tend to eliminate types of proposals involving problem-areas to which they had been alerted. Outside panelists, one might expect, would emphasize selection in terms of substantive topics and, perhaps, methodology.

An inspection of Table 2, shows that in-house screeners tended to screen out 'instruction' as a topic of research (a change of ten percent or more is considered worthy of note in this brief presentation) and screen in research on cognitive processes and that done by psychologists. Otherwise, screeners emphasized methodological characteristics of proposals, favoring those with major strengths in methods and theory, those planning systematic analysis, those likely to meet stated objectives and those that propose complex analysis techniques. It appears that methodological sophistication was an important concern and this emphasis appears to correspond to improvements in analysis sophistication that Travers found among the pre-1973 crop of research findings; an older theme which apparently holds promise of being continued.

Invited panelists shared in-house screener interest in methods strength and tended to favor such projects although the amount of emphasis on methods and on theoretical strength was less than that of screeners. On the other hand, panelists tended to make their impact felt most in substantive areas, throwing out research involving identifiable educational systems, and selecting in topics such as development and attitudes and values. Where proposals involved an educational system, panelist preference went toward systems which were classified as 'established' rather than 'new', 'academic' rather than 'vocational', and involving socialization for 'conformity' rather than 'autonomy'. Research involving behaviors and race/ethnicity as variables was also selected more frequently. Perhaps because of apparent interest in behaviors and development, panelists tended to select in projects involving longitudinal designs. As was true of screeners, panelists tended to favor projects proposing more complex analytic techniques.

In a number of instances, screeners and panelists seemed to have divergent views as can be seen where one group screens out projects with certain features and the other group screens them back in (or vice versa). For example, in-house screeners cut the percentage of projects involving instruction by 16% but panelists increased the percentage by seven. Opposite tendencies occurred for attitudes and values as a topic of study (panelists increased the appearance of this topic), panels tended to favor open schools and screeners favored free schools, but both tended not to favor traditional schools as captured by this typology. None of these latter changes was substantial, however. Finally, panelists seemed more willing than screeners to tolerate projects which were not judged likely to meet stated objectives, perhaps indicating a willingness to take greater risks or a higher priority on substantive topics of interest.

The screening and panel processes had some effect on the distribution of characteristics finally funded---the former more methodological, perhaps, and the latter more substantive---although by and large relative importance of topics remained fairly stable, indicating basic agreement with the distribution of topics originally submitted by principal investigators.

CHARACTERISTICS RELATED TO STRONG DESIGN SCORES

If Travers' assessment of educational research is correct, it would appear that conceptualization and design are more in need of attention currently than analytic or statistical sophistication. We found that the latter appeared to be substantially improved in 1973 and our description of funded proposals for current research suggests that this aspect of research is being given continued attention. How might conceptualization and design be improved? While we have no unique solution to this problem, it is possible to identify in a relatively gross way, some characteristics of these proposals which are associated with higher or lower scores on our overall design scale.

The analytic technique we used for doing this was Automatic Interaction Detection Analysis.[6] This procedure attempts to explain as well as possible the variance on a dependent variable by applying one-way analysis of variance techniques using permutations of a set of categorical independent variables to subdivide the population of proposals into a set of mutually exclusive subgroups. The independent variables employed consist of a number of composite variables developed through cluster analytic techniques and measuring descriptive methodological and descriptive content of proposals, as well as selected academic disciplines of the principal investigators. Each of the 14 predictor variables is measured as a presence-absence dichotomy. The descriptive methodology clusters were: 1)experiment or observation, 2)survey research, 3)common analytic techniques, 4)complex analytic techniques, 5)qualitative research and 6)"other" methods. The descriptive content clusters were: 7)human development, 8)professional training/program/structure, 9)micro-level social systems and 10)macro-level social and political problems. These clusters are briefly explained in the attached note on scales. The selected academic disciplines of the investigators were: 11)education, 12)psychology, 13)sociology and 14)"other" disciplines.

The dependent variable for this analysis was the 20 item research design factor scale which we constructed by applying factor analytic scaling procedures to a larger set of Likert-type rating scales measuring various aspects of design. The scale includes a number of more methodological

features of research but it also includes items such as use of prior research knowledge, discussion of germane literature. explanation of the problem, control of relevant variables and other items which would seem to reflect conceptual strength.

The tree diagram in Table 3 presents the results of this analysis, showing the mean design score for various subgroups of proposals. Examination of the various interactions will be left to the reader, but it should be noted that the highest mean design score for submitted projects occurs among projects which a) propose complex analysis techniques, b) involve experiments or observation as techniques, c) are proposed by principal investigators who are psychologists. At the other extreme, projects with the lowest mean design score are those which: a) do not propose complex analytic techniques, b) use techniques other than experiment or observation, c) do not use survey techniques, d) nor some other defined method, e) do not use qualitative techniques. Although several other variables, notably descriptive content variables, were used in the splitting process, essentially methodological techniques and one category of discipline of principal investigator accounted for the variance in the design scale.

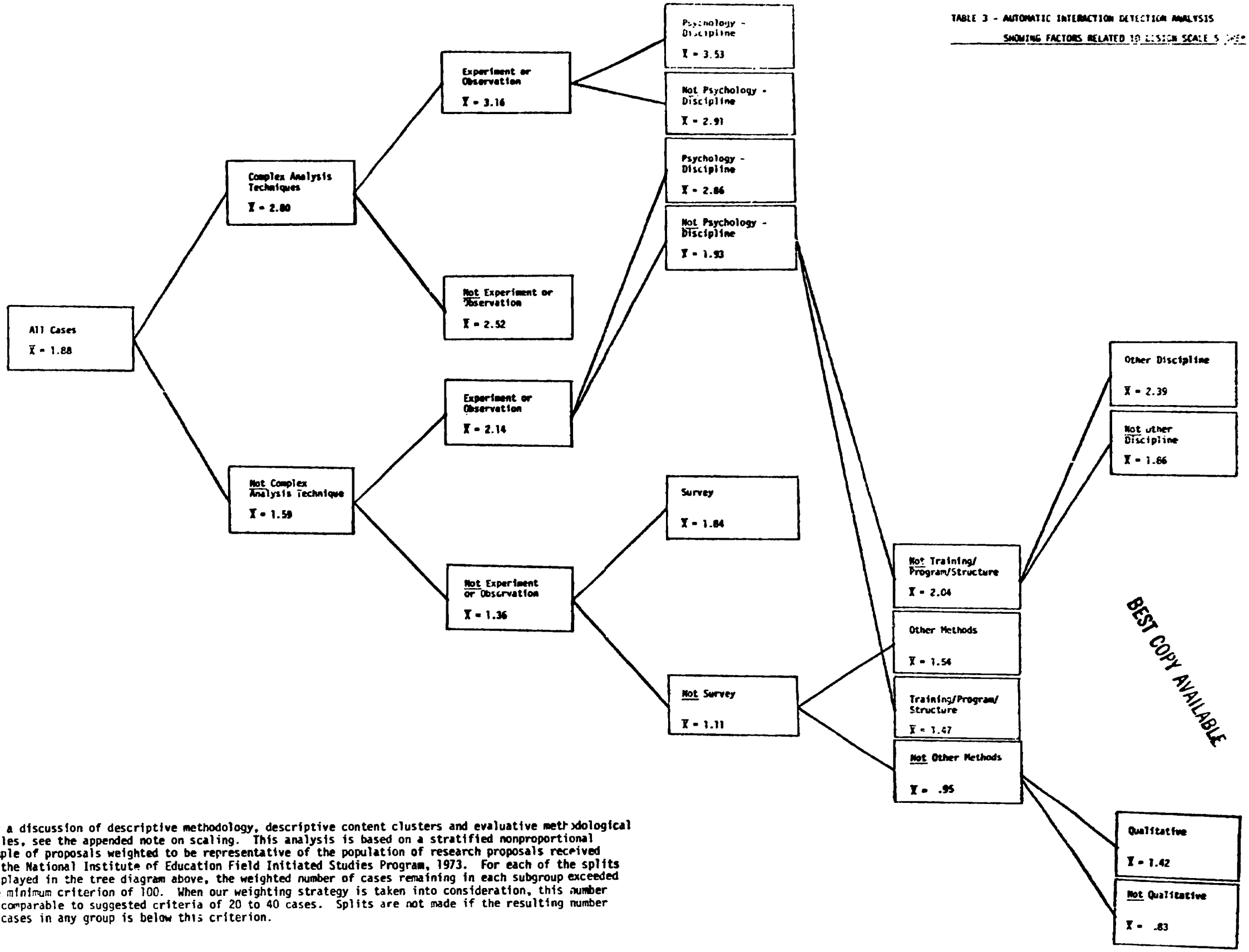
CONCLUSIONS

Our data show that the initial population of proposals submitted to NIE emphasized certain topics rather than others, and that the in-house screeners and the panels expressed to a certain degree their preferences and had some impact in terms of shaping the ultimate characteristics of the research which was finally funded. Further, both the screeners and the panels appeared to be concerned with the conceptualization of the design of research and were able to accomplish improvement in this characteristic of the research funded.

Our objective in this paper was to point out characteristics of proposals which constitute a significant portion of a future cohort of research findings. There are probably many alternative views about the desirability and promise

TABLE 3 - AUTOMATIC INTERACTION DETECTION ANALYSIS

SHOWING FACTORS RELATED TO DESIGN SCALE 5 (MEP



*For a discussion of descriptive methodology, descriptive content clusters and evaluative methodological scales, see the appended note on scaling. This analysis is based on a stratified nonproportional sample of proposals weighted to be representative of the population of research proposals received by the National Institute of Education Field Initiated Studies Program, 1973. For each of the splits displayed in the tree diagram above, the weighted number of cases remaining in each subgroup exceeded the minimum criterion of 100. When our weighting strategy is taken into consideration, this number is comparable to suggested criteria of 20 to 40 cases. Splits are not made if the resulting number of cases in any group is below this criterion.

of this type of research. Finally we hope this description will help focus discussion of educational research objectives by policy-makers, researchers and practitioners.

FOOTNOTES

1. See, for example, Sam Sieber and David Wilder (editors), THE SCHOOL IN SOCIETY, New York, The Free Press, 1973; or Robert M W Travers, editor, SECOND HANDBOOK OF RESEARCH ON TEACHING, Chicago, Rand McNally & Company, 1973.

2. In his introductory comments for the SECOND HANDBOOK OF RESEARCH ON TEACHING (Chicago, Rand McNally, 1973), editor Travers summarized the reactions of contributing authors to the various literatures they reviewed as follows: "In many cases, after reviewing the literature, the author made the decision that the material was such that he could not write a chapter bringing the findings together---the complaint being that the research consisted of a patchwork of unrelated items that neither fitted together nor yielded a useful set of generalizations...although there has certainly been an improvement in the statistical design of studies over the last decade, a corresponding improvement in conceptual design is often not apparent... the heavy emphasis upon what is wrong with educational research...reflects the general level of inadequacy of much of the research that has been undertaken." (Preface, page vii.)

3. FINAL REPORT, METHODOLOGICAL ADEQUACY OF FEDERAL R & D PROJECTS, by Minnesota Systems Research, Inc, supported, in part, by Research Grant SRS 22-P-55936 from the Division of Research and Demonstration Grants, Social and Rehabilitation Service, Department of Health, Education and Welfare, Washington, DC 20201. And ASSESSMENT OF METHODS TO EVALUATE THE SCIENTIFIC RIGOR OF SOCIAL RESEARCH: A STUDY OF SIX REHABILITATION RESEARCH PROJECTS, by Minnesota Systems Research, Inc, sponsored by the Rehabilitation Services Administration and the Office of Research and Demonstrations of the Social and Rehabilitation Service and by the Office of Research and Evaluation Planning of the Assistant Secretary for Planning and Evaluation of the Department of Health, Education and Welfare under Grant Number 22-P-55936.

4. The stability-consistency analyses provided evidence that the raters made stable and consistent judgments. In addition, some comparisons could be made with independently made classifications of projects and these matches were very close. For technical details on measurement issues see NATIONAL INSTITUTE OF EDUCATION FIELD INITIATED STUDIES PROPOSAL REVIEW, 1973, Final Report Contract Number NE-C-00-3-0287 by Minnesota Systems Research, Inc, sponsored by the National Institute of Education, Washington, DC.

5. Travers (op. cit.)

6. J A Sonquist and J N Morgan, THE DETECTION OF INTERACTION EFFECTS, Survey Research Center Monograph #35, Institute for Social Research, University of Michigan, Ann Arbor, 1964.

APPENDIX

A NOTE ON SCALING

Most of the items shown in previous tables are single judgments, but some are scales. We were able to characterize both the substantive topics included in a study and its various methodological features in terms of a large number of separate items selected to cover as broadly as possible the various facets of educational research. These items were used to create three sets of scales which serve to 'locate' a project descriptively and methodologically in an overall fashion. The scales were developed using cluster, smallest space and factor analysis guided by prior conceptualization of content and method. These scales are briefly described below:

DESCRIPTIVE CONTENT: A combination of multidimensional scaling techniques (smallest space, cluster and factor analysis) was used to create four nonmutually exclusive clusters from 21 dichotomous descriptive content variables. The resulting clusters, their component items and the criteria for coding a proposal as being included in a cluster are shown in Table 4.

DESCRIPTIVE METHODOLOGY: Using the same techniques as used for descriptive content variables, 22 dichotomous measures of descriptive methodology were combined into six nonmutually exclusive descriptive methodological clusters. As above, the resulting clusters, their component items and the criteria for coding a proposal as being characterized by the cluster are shown in Table 4.

EVALUATIVE METHODOLOGY

The evaluative methodology scales were constructed through extensive factor analyses of 66 Likert-type rating scales measuring various facets of methodological quality. After determining the factor structure of this set of measures, the scales were "purified" by deleting relatively poor items to maximize their internal consistency reliabilities. Considerable evidence of construct validity was found for the scales (and is discussed in depth in our final report). These scales are described on page 18. The specific items used in measuring each are listed along with internal consistency reliabilities in Table 5.

TABLE 3 - DESCRIPTIVE CONTENT AND DESCRIPTIVE METHODOLOGY CLUSTERS, THEIR COMPONENT VARIABLES & CLASSIFICATION CRITERIA

DESCRIPTIVE CONTENT		
Cluster Name*	Variables**	Criteria***
Human Development	Achievement Instruction Ability & Aptitude Children Cognitive Processes Individual Development Intra-Individual Level of Analysis	3
Professional Training/ Program/ Structure	Preparation for Professional Personnel Program School System Structure	2
Micro-Level Social Systems	Classroom as a Social System Classroom Role Structures Internal Dynamics of a Small Group	2
Macro-Level Social and Political Issues	Civil Rights Community Culture Equality of Educational Opportunity Discrimination Educational Policy Educational Governance Institutional Complex and Linkages	2

DESCRIPTIVE METHODOLOGY		
Cluster Name*	Variables**	Criteria***
Experiment or Observation	Laboratory Experiment Observation	2
Survey Research	Statistical Sample Survey Cross-sectional Study Questionnaire or Test Data	2
Common Analytic Techniques	Univariate Description Percentaged Tables Cross-Tabulations	2
Complex Analytic Techniques	Computerization Measures of Association Correlation Inferential Statistics Analysis of Variance Confidence Limits Multiple Regression Factor Analysis	3
"Other" Methods	Other Type of Research Other Source of Data	2
Qualitative Research	Informal Interviews Interview Qualitative Analysis Content Analysis	2

*These non-mutually exclusive clusters were constructed through separate factor, cluster and smallest space analyses, of which an explanation appears in our final research report.

**Each of these variables was measured as a "presence-absence" dichotomy. All are non-mutually exclusive.

***Number of variables needed to be characterized by a cluster.

1. Design implies those activities which involve planning, rather than conduct of research or analysis of the data. The most critical features of research design include formulation of the problem, familiarity with prior relevant work, awareness of basic assumptions, definition of key concepts, statement of hypotheses, specification and measurement of key variables and control of other relevant variables.

2. The measurement validity factor concerns only the reliability and validity of measures. Its focus is narrow in that it involves neither operationalization and quantification per se nor issues of validity surrounding representativeness of sampling.

3. The sampling factor refers to the selection of one or more cases to study from some larger class of potential cases. This includes definition of the target population, appropriateness of the sampling unit and sample size, and representativeness of the sample.

4. The use of statistical procedures factor concerns the analysis stage of research. Specifically, it includes a range of quantitative statistical procedures and techniques. Its limitations involve its exclusion of non-statistical qualitative analytic procedures, of some very ordinary quantitative analytic procedures (such as univariate description, percentage tabling and crosstabulation), and of items centering around quality of analysis and interpretation of findings.

5. The care of analysis factor includes items focusing on the accuracy, orderliness and thoroughness of the analysis and the accuracy and appropriateness of the conclusions. It does not include items focusing on the care of design, sampling or measurement.

TABLE 5 - EVALUATIVE METHODOLOGY FACTORS AND THEIR COMPONENT VARIABLES

Factor*	Cronbach's Alpha	Variables
Design**	.95	Description of main concepts Explanation of research design details Description of design in application Specification of dependent variables Specification of independent variables Specification of control variables Discussion of germane literature Use of available prior knowledge Specification of hypotheses Discussion of assumptions Manageability of problem Experimental design Provision for hypothesis generating experiences Use of existing measurement procedures Quantification of variables Provision for adequate research resources Design efficiency Explanation of the problem Fit of design to the problem Control of relevant variables
Measurement Validity**	.83	Use of scaling procedures Pretest of procedures and instruments Handling of validity issues Handling of reliability issues Checks on validity of instruments
Sampling**	.79	Definition of the population of interest Use of probability sampling procedures Appropriateness of sampling unit Appropriateness of sample size Random selection of cases
Statistical Procedures***	.82	Computerization of the analysis Use of measures of association Use of inferential statistics Use of correlational analysis Use of multiple regression analysis Use of analysis of variance Use of statistical controls Use of factor analysis
Care of Analysis***	.76	Accuracy, orderliness and care of analysis Logic of data-conclusions linkage Data support for conclusions Completeness of analysis of available data Examination for possible biases

*Factor scores were computed by giving unit weightings to each of the component items for each factor, by summing the ratings, and by dividing by the number of component items. Scores on the first 3 factors have a potential range of 0-5, while scores on the last 2 factors can potentially range from 0-1.

**The rating categories of the variables comprising these factors ranged from 1 (unacceptable) to 5 (far above average), with 0 indicating insufficient information or no consideration being given to a variable.

***The variables comprising these factors were judged as present (1) or absent (0) in a given proposal.