May 1972

A Design for Advanced Preparation of Educational Researchers. (Volumes I and II)

Fon R. Belcher
East Tennessee State University

A DESIGN FOR ADVANCED PREPARATION OF EDUCATIONAL RESEARCHERS

An Abstract

Presented to

the Graduate Faculty

East Tennessee State University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

---

by

Fon R. Belcher

May 1972

Fon R. Belcher, B. S., Morehead State University, June, 1964.
M. A., East Tennessee State University, June, 1970.
Ed. D., East Tennessee State University, May, 1972.

A DESIGN FOR ADVANCED PREPARATION OF EDUCATIONAL RESEARCHERS

Statement of the problem. The central problem of this study was twofold: (1) to determine whether the educational research preparation offered by selected colleges and universities bore out the findings of critics, and (2) to synthesize and synergize opinions and facts regarding educational research, as expressed by men active in the field, into a flexible proposal whereby individual student's programs in research preparation could be "custom-made" to suit their most obvious research needs.

Summary of assumptions. It was assumed that: (1) existing university research programs for the preparation of school administrators and educational research personnel reflect the philosophy, beliefs, and work-styles of those who plan and operate them. (2) A study of the common elements between practices of trainers, opinions of professors and students of educational research, and the findings of research provide the best data to consider prior to the design of a new advanced research preparation program. (3) A program could be developed which would be superior in content and student experiences to any single program available in the selected institutions.

Procedures. On-site visits were made to universities selected for the sample for the study. Discussions with personnel associated with educational research (instructors, department chairmen, students, others) were conducted using the interview technique and a response data form. All interviews were tape recorded. A survey and analysis of literature were conducted to ascertain the most widely discussed practices in the research training of advanced graduate students in education. Interviews with prominent educational research personnel and doctoral candidates associated with the selected institutions were conducted. All of the data thus collected and analyzed were considered as nuclei in the proposed research program design resultant from the total procedure.

Summary of findings, conclusions, and implications. Each of the research authorities described a plan or procedure of a research preparation program but did not prescribe a sequence of courses for an adequate program. Each held to the traditional research methods,

statistics and measurement courses offered by all universities having doctoral programs. Yet, in each instance, they strongly recommended that a proper research preparation program could and should be developed wherein the student actually worked as an intern to someone engaged in educational research.

A major conclusion of this study relates to Byar's statement concerning the "bedevilment theory"; that one cannot get the man out of his research. All those involved in this study revealed a personal approach to educational research and research training. That is, any course is like the professor who teaches it. Also concluded was that research preparation for the practitioner must be "custom-made" for competency in consuming research and developing decision-making skills.

The dilemma of research preparation programs, the constituted a major implication of this study. That is, each student has his own unique "work-style." His program should be "custom-made."

. . . .

Dissertation prepared under the guidance of Dr. William L. Evernden, Dr. Madison Byar, Dr. Howard Bowers, Dr. Harold Measel, Dr. Charles Burkett, and Dr. Henry Swann.

ACKNOWLEDGMENTS AND DEDICATION

ACKNOWLEDGMENTS

Throughout this dissertation, emphasis is placed on the role of the student's advisory committee, through its chairman. Working with and through Dr. William L. Evernden, this point of view became an active ingredient in my own advanced program. His fine advice and his behavior as my chairman constantly reassured me of the student benefit to be accomplished through the theme of my dissertation.

DEDICATION

This dissertation is dedicated to T. Madison Byar, whose efforts in the development and implementation of The Advanced Graduate Program at East Tennessee State University has made him a "significant other" to me.

F. R. B.

Volume I

A REPORT OF THE RESEARCH ACTIVITY
ACCRUENT FROM THE DESIGN
OF THIS STUDY

Volume II

A PROPOSED DESIGN FOR ADVANCED
PREPARATION PROGRAMS IN
EDUCATIONAL RESEARCH

TABLE OF CONTENTS

Volume I

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Once in the last third of the twentieth century, educators became disillusioned with the lack of adequate evaluation and accountability in the discipline of education. In 1965 in his book, A Cross-Section of Educational Research, Edwin Wandt noted and discussed the sad state of educational research. His discussion centered on the utter lack of adequately reported research which appeared in the journals used in his study. Several judges or evaluators, who were research writers themselves, critiqued articles dealing with research in educational journals. One of their first handicaps was the difficulty of finding articles that could be classified as true research. Journals for the preceding five years were searched and a panel of judges analyzed each research article on its merits. Articles were rated on an instrument devised for evaluating studies. The findings, as rated on the instrument, showed that the research that appeared in educational journals was average to mediocre in quality. Wandt pointed out the necessity for educators to develop improved research ability if they were to publish creditable results. If education was to be a profession it demanded that scholars be prepared to conduct experiments which could be validated and the results of which would stand up under other forms of professional evaluation.

Since education was not an exact science it was difficult to measure the results of experimentation. However, it was not impossible for educators to familiarize themselves with the various methods of research including statistical procedures, and computer applications to educational problems.[1]

Educational research, as a part of the general discipline of education, was seen to be composed of several categories of research-oriented activity: basic research, the creation of new knowledge; applied research, the field testing of basic research findings, and the action type of research which was seen more as a linguistic activity than as either a search for new knowledge or a testable application of research findings. Beyond these aspects there seemed to be the practitioners of the profession of education who, so to speak, "consumed" research; that is, they fell into established patterns of activity resultant from research efforts other than their own.

These facts and conditions pointed to the deficiencies in extant preparation programs in those colleges and universities that profess to offer programs to prepare educational researchers. There were, in 1972, programs so mathematically centered that surveys, historical research, time studies, comparative analysis and studies of a modern existential nature were not considered research. At other institutions the flexibility of research preparation programs approached the "no-orientation" level. The research carried out in the preparation of this study became the backdrop for the design of a "desirable" program that would break

---

[1]Edwin Wandt, A Cross-Section of Educational Research (New York: David McKay Company, Inc., 1965), passim.

the total field of educational research into phases or components suited to the work-styles of prospective researchers.

THE PROBLEM

Statement of the problem. The central problem of this study was twofold: (1) to determine whether the educational research preparation offered by selected colleges and universities bore out the findings of critics, and (2) to synthesize and synergize opinions and facts regarding educational research, as expressed by men active in the field, into a flexible proposal whereby individual student's programs in research preparation could be "custom-made" to suit their most obvious research needs.

Conclusions were drawn from a comparative analysis of programs at the University of Cincinnati, Miami (Ohio) University, the University of Tennessee, and the University of North Carolina through on-site visits and from the opinions of personnel involved in educational research at The Ohio State University, the University of Tennessee, the University of Cincinnati, and the University of Oregon through recorded telephone interviews.

The subproblems treated herein are as follows:

a. A design for researchers.
   How much and what kinds of research training and experiences are needed for those who are going to participate in basic research?

b. A design for practitioners (school administrators).
   How much and what kinds of research training and experiences are needed for those who are going to limit themselves to action research and the application of research findings?

Importance of the study. The research carried out in the preparation of this study is principally important because it gives the reader an opportunity to evaluate certain university research programs on the basis of their sponsors' opinions and in terms of the designs of their programs. Such an analysis of designs of existing programs leads naturally into the major purpose of this study; namely, the development and presentation of ideal or custom-made programs of courses and experiences appropriate to the advanced training of educational researchers. Progress in the applications of technology to research outside the area of professional education is a significant and continuing consideration in designing advanced training programs; they must be up-dated as further technological advances are made.

This study represents an attempt to use the knowledge gained from an analysis of existing programs in the development of a program which will include the best that is now known about the advanced training of educational researchers.

Significance of the study. The significance of this study derives from the fact, established by the interviews carried out and by an analysis of institutional programs in the area of educational research, that the type of research preparation a student needs is determined by the kinds of research one may be called upon--or may call upon himself--to carry out. Until a problem, calling for research activity to solve it, exists and is identified, no student needs research preparation. However, beset as all of education is with a variety of problems, a further significance of this study lies in the explanation it offers for why more educators do not feel competent to attack problems through the applications of research techniques.

The beginning, in 1971, of advanced graduate programs in educational administration and supervision at East Tennessee State University caused both the professors and the initial group of students to consider the problem of adequate preparation in the areas of educational research for advanced students.

A study of preparation programs in similar institutions was conceded to be a good place to get an estimate of "what was going on." Beyond this, the study became locally significant because from the necessary survey, with its series of interviews with professors and students, ideas for planning a research preparation program for East Tennessee State University were gathered, analyzed, and woven into a program proposal.

The proposal resultant from the above research reflects the variety of types and degrees of research preparation needed for the functional preparation of advanced students whose work-styles and ambitions are varied. It was anticipated that the outcomes of this study would be helpful to advanced graduate students and professors of educational research, and would have applicability to advanced graduate programs anywhere in the nation. That fact alone adds a significance to the study beyond that attached to many of the "standard" programs studied.

DEFINITIONS OF TERMS

Analysis. Analysis is breaking down, ordering, summarizing, and testing data so that answers to research questions can be obtained.[2]

Application. Application means the use of the computer in dealing with research data analysis techniques.

Computer. Computer is an 1130 International Business Machines disc system with a thirty-two storage capacity. Alternate disc packs provide for unlimited storage of data. A computer contains five components: an input system, storage or memory, a control system, an arithmetic unit, and an output system.[3]

Design. A design is a plan or pattern for theorizing. It includes organizing, analyzing, and interpreting data collected in a research study.

OPPO. OPPO refers to the way research and research training is Organized, Planned, Programmed, and Operated.

Practitioner. A practitioner is a producer of decision research and a consumer of research who implements knowledge in a particular context.[4]

---

[2]Paul Monroe, Encyclopedia of Educational Research (New York: The Macmillan Company, 1969), p. 1138.

[3]Gilbert Sax, Empirical Foundation of Research (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), p. 408.

[4]John H. M. Andrews, "Differentiated Research Training for Students of Administration," Educational Research: New Perspectives, eds. Jack A. Culbertson and Stephen P. Hencley (Danville, Illinois: The Interstate Printers and Publishers, 1963), pp. 359-61.

Research. Research is an approach to an understanding of the universe along a broad thoroughfare of organized knowledge solidly established on observation and experiment imbedded in a matrix of theory.[5]

Researcher. A researcher is one who participates in basic research for the achievement of a fuller knowledge or understanding of the subject matter under study, rather than making practical applications of new knowledge.[6]

Statistics. Statistics is a technique for summarizing numerical data and comparing obtained results to chance expectations.[7]

Technique. Technique is any organized procedure for gathering and/or analyzing research data.

Tool skills. Tool skills is an understanding of research implements in computational, quantitative, and information sciences in the fields of education, mathematics, and psychology.[8]

Work-style. An individual's work-style is determined by his beliefs--by those operational factors in his thinking that determine

---

[5]Palmer O. Johnson, "Introductory Remarks at Opening of the Symposium on Educational Research," First Annual Phi Delta Kappa Symposium on Educational Research (Bloomington, Indiana: Phi Delta Kappa, 1960), p. 15.

[6]Andrews, op. cit., p. 358.

[7]Monroe, loc. cit.

[8]Clifford James Houston, "Trends in Research Tool Requirements for the Doctorate in Education" (unpublished Ph.D. dissertation, Colorado State College, 1969), passim.

both what type of things he is most likely to do and the way or ways he is most likely to go about doing them.[9]

ASSUMPTIONS

It was assumed that:

(1) Existing university research programs for the preparation of school administrators and educational research personnel reflect the philosophy, beliefs, and work-styles of those who plan and operate them.

(2) The universities studied require research course sequences of advanced graduate students, supply those students with bibliographical lists of research books and references, and use textbooks prescribed by the professors.

(3) In the universities studied, computer centers have selected programs accessible to research students who have "hands-on" access to computers.

(4) Professors actively engaged in research could help identify the order of priority of research courses and experiences.

(5) A study of the common elements between practices of trainers, opinions of professors and students of educational research, and the findings of research provide the best data to consider prior to the design of a new advanced research preparation program.

(6) Advanced graduate students planning careers in educational research or large scale institutional planning would need to complete a research program of greater depth and breadth than students who anticipate other educational careers.

---

[9]Statement by T. Madison Byar, personal interview, January 20, 1972. (Tape recorded.)

(7) Most school administrators do not usually have time to perform most types of research, and need only to understand research and to develop the ability to interpret and use research results; they will, therefore, employ experts to do research for them.

(8) A program could be developed which would be superior in content and student experiences to any single program available in the selected institutions.

LIMITATIONS

This study was limited to an analysis of printed programs, the opinions of professional research personnel, and suggestions made by graduate students of educational research and analysis. Supportive data were gathered from university programs and interviews with personnel associated with the University of Tennessee, Miami (Ohio) University, the University of Cincinnati, the University of North Carolina, The Ohio State University, and the University of Oregon.

Data were obtained from recorded phone conversations, face-to-face recorded conversations with nationally recognized authorities in educational research, practicing school administrators, and research and development planners at the universities of Oregon, Tennessee, Miami (Ohio), Cincinnati, East Tennessee State, North Carolina, and (The) Ohio State. Some interviews were held with university professors and other personnel, prepared in research methodology but not currently engaged in research. This latter group discussed ideas related to research and research preparation programs and voiced opinions developed principally in graduate schools. Other sources of data included papers, brochures, and research proposals appropriate to the possible solution of the problem of the study.

PROCEDURES

Universities with recognized research programs and research-competent personnel were selected to serve as the sample for this study. Interview Response Data Forms (Appendix A) were used to obtain basic data concerning the characteristics of programs, and the personnel who man them in each of the universities.

On-site visits were made to four of the universities listed. Discussions with the instructors of research courses provided information on course content, student participation and follow-up, and text and reference materials. Previously prepared data-gathering instruments were completed during the visits. Analyses were made of what the computer centers had to offer research students. Relevant data from each of the universities were gathered, organized, and analyzed. From this analysis, the most effective facets of each program were identified and described in tables and in narrative.

Telephone interviews were recorded with research personnel at two universities not visited, and with people not available during the visits to one other university. Answers to appropriate questions from the Interview Response Data Forms and other questions specific to the role of the interviewee were recorded and used as data supplemental to that gathered during on-site visits.

A survey and analysis of literature were conducted to ascertain the most widely discussed practices in the research training of advanced graduate students in education.

Interviews with prominent educational research personnel and doctoral candidates associated with the selected institutions were conducted through the use of an interview format included in Appendix A.

A semantic differential form (Appendix B) was administered to students in educational research classes at East Tennessee State University. The attitudes of these students were factored out of the data for use in the development of program designs.

All of the data thus collected and analyzed were considered as nuclei in the proposed research program design resultant from the total procedure.

ORGANIZATION OF THE STUDY

This study was organized so that it could be reported in two volumes. Volume I deals with the implementation and reporting of data resultant from carrying out the research design. Volume II deals with a proposed research preparation program for advanced students in the discipline of education.

Volume I, Chapter 1 includes an introduction to the problem, subproblems, importance of the study, significance of the study, assumptions, definitions of terms, and limitations of the study. The procedure for conducting the study is outlined in this chapter.

Volume I, Chapter 2 contains a review of the related literature dealing with educational research.

Volume I, Chapter 3 reports the findings of the study.

Volume I, Chapter 4 is devoted to a brief outline of suggested programs for the educational research consumer and for the educational researcher.

Volume I, Chapter 5 presents the summary, conclusions, and implications of the research findings.

Volume II, Chapter 1 presents a program for an introductory course in research methods and experiences, including design and techniques.

Volume II, Chapter 2 presents a plan for an advanced research course for the educational administrator.

Volume II, Chapter 3 contains suggestions for a seminar in writing and interpreting research reports.

Volume II, Chapter 4 outlines a program for educational research as a major field of study.

Chapter 2

REVIEW OF RELATED LITERATURE AND RESEARCH

Writers in the past, who have described research training programs, have generally dealt only with methodological content. Sieber sampled sixty-four graduate schools of education; twenty-five indicated that they neither emphasized research training nor provided a program for training researchers in education. Seventeen percent emphasized research training and provided some form of program for students who wanted to make research a career. Sieber noted that only 27 percent of the responding universities offered majors in educational research. Basic Research Methods and Design, a course most frequently offered by schools, was offered by 96 percent of the schools. There was an average of 4.6 courses per school. Statistics was offered by 85 percent of the schools with an average of 3.0 courses per school. Testing and measurement was offered by 71 percent of the schools, with 2.0 courses per school. Courses based on current research activities were offered in 22 percent of the schools, with 2.2 courses per school. Courses on School Surveys were offered by 16 percent of the schools, with 1.5 courses per school. (Other unspecified research courses were offered in 14 percent of the schools, with 2.7 courses per school.)[1]

---

[1]Sam D. Sieber, Analysis of USOE Research Training Programs, Bureau of Applied Social Research (New York: Columbia University, 1968), p. 10.

Krathwohl made a survey of doctoral producing universities. Ninety-one percent of the institutions offered introductory courses in research, a greater percentage offered statistics; 46 percent offered experimental design courses; and courses in measurement and evaluation were frequent. However, the full sequence of courses--research methods, statistics, and design and measurement, as a requirement in doctoral programs, was quite rare.[2]

Roaden searched for new or different programs. He received program descriptions, adequate for analysis, from forty-seven institutions. These institutions described fifty-seven programs: two post-doctoral, thirty-one doctoral, five masters, six undergraduate, and thirteen short-term institutes.[3] Roaden concluded that the evidence indicated that there were still some unsolved problems which were identified in Clark's 1963 observation that "educational research at this point in its historical development was clearly inhabiting the periphery of the profession."[4]

Sproull studied the relation of research preparation to duration, attrition, and research communication of doctoral candidates

---

[2] David R. Krathwohl, "Current Formal Patterns of Educating Empirically Oriented Researchers and Methodologists," The Training and Nurturing of Educational Researchers (Bloomington, Indiana: Phi Delta Kappa, 1965), passim.

[3] Arless L. Roaden, "An Analysis of Formal Programs for the Training of Educational Research and Dissemination Personnel," A Study of Roles for Researchers in Education, eds. David L. Clark and John E. Hopkins (Bloomington, Indiana: School of Education, Indiana University, 1969), passim.

[4] David L. Clark, "Educational Research; a National Perspective," Educational Research: New Perspectives, eds. Jack A. Culbertson and Stephen P. Hencley (Danville, Illinois: The Interstate Printers and Publishers, Inc., 1963), pp. 33-42.

in education at Michigan State University. The average doctoral student completed 5.5 courses categorized as research preparation. The average number of courses in each kind of research preparation was approximately two courses in theory, theory construction or logic; one course in research methods; one in measurement or evaluation; and 1.5 courses in statistics and mathematics. The average number of research preparation courses ranged from approximately three to eight. Candidates from the Departments of Counselling, Personnel Services, Educational Psychology, and candidates with cognates in psychology received their major research preparation in theory, measurements and statistics. Candidates from the Department of Secondary Education and Curriculum and candidates with cognates in sociology were prepared chiefly in research methods.[5]

Millikan reported on eleven of the most important characteristics of institutions that were high producers of researchers. Of these characteristics, five related to the apprenticeship mechanism, four to the formal mechanism, and two to the peer group mechanism.[6]

Houston studied trends in research tool requirements for the doctorate in education. His study showed that institutions granting the doctorate indicated that there was usually one research tool

---

[5]Natalie Loraine Sproull, "The Relationship of Research Preparation to Duration, Attrition and Research Communication of Doctoral Candidates in Education (unpublished Ph.D. dissertation, Michigan State University, 1969), passim.

[6]Nancy Horne Millikan, "Processes of Socialization in Educational Research: Delineation of Problems Facing the Development of Researchers by Graduate Institutions of Education" (unpublished Ph.D. dissertation, Columbia University, 1969), passim.

requirement for the Doctor of Education degree, while in most institutions granting the Doctor of Philosophy degree in Education, there were typically two tools of research required. The fact that the educator of the future needs to be a competent research worker demands a careful examination and analysis of the present trend in research tool requirements for the doctorate in education.[7]

Based upon an analysis of these data, Houston concluded that institutions which granted the doctorate in education differed from each other relative to research tool requirements, methods, and foreign language, when these institutions were categorized according to type of institution and total enrollment. Forty-seven percent of those institutions granting the Ed.D. required no research tool for that degree, while the majority of institutions that granted the Ph.D. in Education required two research tools for that degree. The majority of institutions granting the Ed.D. allowed course work in applied statistics to satisfy a research tool requirement. The majority of institutions granting the Ph.D. in Education allowed the Graduate School Foreign Language Test of the Educational Testing Service, or a foreign language department test, to satisfy their research tool requirement. The majority of institutions which allowed a foreign language to satisfy a research tool requirement reported that German, French, Russian, and Spanish were acceptable languages for this purpose. A majority of institutions granting the Ed.D. allowed course work in applied statistics to satisfy the research tool requirement because they were convinced of

---

[7]Clifford James Houston, "Trends in Research Tool Requirements for the Doctorate in Education" (unpublished Ph.D. dissertation, Colorado State College, 1969), passim.

its utility. A majority of the institutions granting the Ph.D. in Education allowed course work in applied statistics to substitute for one language requirement. Perhaps an emphasis on scientific technology and corresponding advances in computer science were partially responsible for the relaxation of the language requirement. It was recommended that all institutions granting the doctorate in education immediately instigate follow-up studies of their doctoral students to ascertain which research tool requirements and methods were most helpful to their students after graduation, and that they modify their research tool requirements accordingly.[8]

Altschuld studied the measurement and analysis of factors predictive of graduate student success in educational research. The Research Orientation Index, a sixty-item Likert type scale, and The Research Knowledge Index, a fifty-item multiple choice research competency test, was constructed. Based upon a review of literature, the following variables were used to predict graduate student success in the program: research knowledge, prior research experience, sex, research orientation, interest in course work, interest in activities, age, and degree sought. The results indicated that graduate student success in research was predictable and would be even more so with more refined instruments.[9]

---

[8] Ibid.

[9] James William Altschuld, "A Study of an Experimental Training Program in Educational Research and Development: The Measurement and Analysis of Factors Predictive of Graduate Student Success" (unpublished Ph.D. dissertation, The Ohio State University, 1970), passim.

The American Psychological Association report summarized the view that the best preparation for research was apprenticeship to a skilled researcher. Course work, formal examination requirements, and anything else that could be standardized, constituted what was ancillary to research training.

> What is of the essence is getting the student into a research environment and having him do research with the criticism, advice, and encouragement of others who suffer the same pain and enjoy the same rewards. Research is learned by doing and taught mainly by contagion--Research must first be going on if there is to be research training.[10]

Richard Judson Puffer conducted a survey of the 727 members of the American Association of Colleges for Teacher Education (AACTE) to determine their involvement with and capabilities for educational research. Puffer recommended that institutions develop training programs to serve as springboards for federal efforts, research institutes, and workshops.[11]

Buswell and McConnell defined several roles of the educational researcher, including the educational development specialist. His role behavior involved the design of educational practices using supporting materials and equipment drawn both from available general knowledge (basic research) and knowledge about user demands and requirements. A sub-specialization of educational development specialists was the retriever-converter whose role behaviors were to scan the literature of

---

[10]American Psychological Association, "Report of the Seminar on Education for Research in Psychology," American Psychologist, XIV (April, 1958), 167-69.

[11]Richard Judson Puffer, "The Educational Research Involvement and Capabilities of Institutions for Teacher Education" (unpublished Ph.D. dissertation, Northwestern University, 1967), passim.

research and practice, and codify it in a fashion permitting easy access for the purpose of translating or converting knowledge into workable prototype programs in schools and universities. The field tester as a researcher exercised major responsibility for assessing the workability, consequences, and feasibility of a particular educational innovation, usually in a preliminary or pilot stage. The quality control man's chief role behaviors centered around the routine assessment of the consequences of educational practices, once installed and in regular use. The change agent's catalyst's needed role behaviors included aiding school system occupants with strategy planning and the installation of change-planning mechanisms, and the design of needed in-service programs at the diffusion/implementation stage. Such change agents may be externally located as sociological strangers or directly employed by and sited in local schools. The county agent was a researcher. Though agriculture was an imperfect model for education, it seemed likely that a suitably transformed "county agent" role would be useful. The role behaviors were those of interpreting research findings to potential user groups. Utilization specialists must have a substantial part of their training in university settings. Internships and intensive field work seemed crucial in roles which served to link developmental functions in the knowledge flow.

Researchers in the past have not been competent in role behaviors requiring the construction of data-collecting tools for practitioners' use, situation-focus information retrieval, diagnostic research and feedback, self-study facilitation or training for research utilization and facilitation of diffusion in practice.

Training programs for such role behaviors need to include not only some involvement with field research and research utilization sites, but some clear conceptualization of research utilization processes in educational systems. If universities ignore this, increasing disjunction between knowledge-producers and users may be expected, without the likelihood of coherent development in either educational practice or theory.[12]

L. Craig Wilson, T. Madison Byar, Arthur S. Shapiro, and Shirley H. Schell discussed the predicament of the supervisor practitioner when faced with research possibilities.

> His usual response is to encourage others verbally to undertake studies or experiments, but studiously avoid his own personal involvement. Another is to assume the role of research publicizer, diffuser, applier, and interpreter, but again to dodge a direct commitment to become a producer of research. Still another response is to authorize externally conducted and directed studies and surveys which use local school subjects, but demand no other form of participation, including responsibility to implement recommended changes. Finally, there is his option of rejecting research altogether, or of redefining the term to make the entire process more compatible with his own perceived reality even though the classicists consider it to be a kind of non-research. The latter alternative is the one that seems to be gaining the most momentum.[13]

Julian Stanley proposed that school systems release (for fifteen months) a person with a liberal arts background and interest in research. This person should apply to universities for a special fellowship to be

---

[12]Guy T. Buswell and T. R. McConnell, Training for Educational Research, Cooperative Research Project No. 51074 (Berkeley: University of California, 1966), passim.

[13]L. Craig Wilson, T. Madison Byar, Arthur S. Shapiro, and Shirley H. Schell, Sociology of Supervision, An Approach to Comprehensive Planning in Education (Boston: Allyn and Bacon, Inc., 1969), p. 335.

funded partly by the USOE and other organizations and partly by the school system. In return, the person would agree to work at least two years as an educational research specialist in the school system that supplemented the fellowship. This person would qualify for a Master's degree in Educational Research.[14]

Fleury studied the facets relevant to the development of applied educational research training programs. He concluded that research training itself was not on a firm research foundation. There was a paucity of research on researcher training, with no comprehensive experimental studies published to date, and only four descriptive studies with a fifth study still in progress.[15]

The projected manpower shortage by 1974 of nearly 10,500 professional research and development personnel was the context in which a rationale was presented for the development of a training program for researchers at East Tennessee State University.[16]

As stated in the introduction to the problem, educational research was severely criticized for its inadequacies. Educators'

---

[14]Julian Stanley, "Preparing Educational Research Specialists for School Systems: A Proposal," Phi Delta Kappan, XLVIII (December, 1966), 110-13.

[15]Bernard J. Fleury, Jr., "A Study of Factors Relevant to the Development of Applied Educational Research Training Programs" (unpublished Ed.D. dissertation, University of Massachusetts, 1968), passim.

[16]Joseph Suyeo Sakumura, "An Analysis of an Experimental Research and Development Program for Talented Undergraduate Education Students" (unpublished Ph.D. dissertation, The Ohio State University, 1969), passim.

indifference to research was mentioned by Travers[17] and Coladarci[18] as a factor contributing to these inadequacies. Fattu[19] found that only ten of ninety-four institutions encouraged research. He surveyed ninety-four institutions which offered doctoral programs. The indifference was attributed in part to the practitioners' domination of the institutions.

The single most important technological factor in changing educational research in the past decade was the computer. Computerization of data analysis techniques was influential. The computer made lengthy and complex analyses possible. The nature of research problems changed because of the influence of the computer. The computer had side effects; it was a teacher because it forced the researcher to learn his methods in depth. Also, the researcher needed to understand computer usage and programming. The computer forced significant changes in graduate programs.

RELATION OF PROPOSED STUDY TO PREVIOUS RESEARCH

A review of the literature of Sieber, Krathwohl, Houston, Sproull, Coladarci, and Travers indicated that offerings of some type of preparation in methods of research for doctoral students was included by the majority of universities they studied. Educational research

---

[17]Robert M. W. Travers, An Introduction to Educational Research (New York: The Macmillan Company, 1958), passim.

[18]Arthur P. Coladarci, "Toward More Rigorous Educational Research," Harvard Education Review, XXX (1960), 3-11.

[19]Nicholas A. Fattu, "A Survey of Educational Research at Selected Universities," First Annual Phi Delta Kappa Symposium on Educational Research (Bloomington, Indiana: Phi Delta Kappa, 1960), pp. 1-21.

preparation programs consisted of sequences of courses, usually Methods of Research, Introduction to Statistics, Design, and Measurement. Although these "courses" were offered by the universities they studied, the complete sequence of course work was available but not required in most instances.

This paper begins with a review of literature related to research preparation programs and develops into a design for preparing educational researchers. The major difference in this proposal as opposed to the programs and authorities discussed in the review of literature is that the design includes the opinions of people active in educational research. Byar described the "bedevilment of education" as "unlike biology or physics, you cannot get the man out of his research. The results will be tempered by his philosophy."[20]

This proposal cuts across the total field of educational research preparation programs and introduces the idea of "prospective work-styles" of probable researchers. This relationship is further amplified by an analysis of interviews conducted with leading authorities in educational research, as described in Chapter 4 of this paper.

SUMMARY

Figure 1 illustrates the dilemma of thinkers and writers who deal with the problems of educational research. There definitely is an area of common agreement about the tool areas; but, in the possible

---

[20]Statement by T. Madison Byar, personal interview, January 20, 1972. (Tape recorded.)

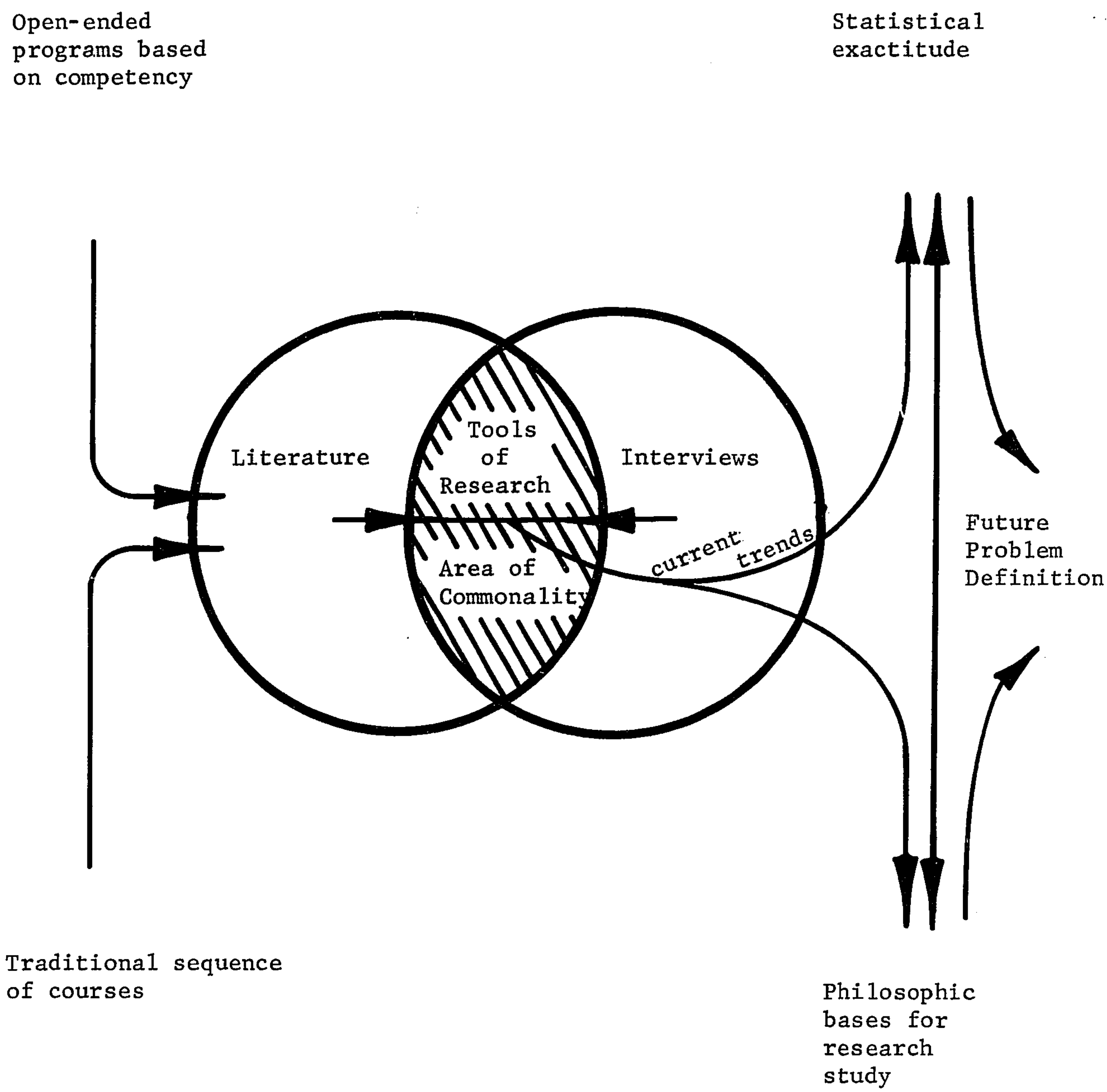

Figure 1

The Dilemma of Planning Educational Research Programs need for and uses of those research tools there are vast differences of opinion ranging all the way from that of Frank Baker who stated, "I don't think that education has in fact a knowledge base of its own. Our knowledge base comes from every field except education and all those roads and all those areas lead to the mathematics department,"[21] to the opinion of Ralph Purdy that "basic beliefs and values tend to control the research that is carried on." It is "the man, the problem, and the ultimate result,"[22] all of which implies what is illustrated by the white areas of Figure 1, which represent the problems of this study.

---

[21]Julian Stanley (ed.), Improving Experimental Design and Statistical Analysis, citing Frank Baker in a discussion on Dr. Stanley's paper, "On Improving Certain Aspects of Educational Experimentation," at the Seventh Annual Phi Delta Kappa Symposium on Educational Research (Chicago: Rand McNally and Company, 1967), pp. 45-46.

[22]Statement by Ralph Purdy, personal interview, December 1, 1971. (Tape recorded.)

Chapter 3

ANALYSIS OF INTERVIEWS AND ON-SITE VISITS

The following is an analysis of information obtained from individuals associated with the teaching of research and the development of research programs in four of the universities included in the study.

Table 1 indicates that the number of instructors teaching research courses was not significant where enrollments were small. The usual effort was to meet the minimum graduate requirements of the school. Tables 1 and 2 viewed together indicate that in Universities B and C the instructional task relative to educational research was probably a minor part of the instructor's total responsibility. Universities A and D provided opportunities for instructor specialization in educational research.

Table 1

Number of Instructors Who Regularly Teach Educational Research Courses

University	Instructors
A	2
B	4
C	3
D	3

Minimum post-baccalaureate requirements in educational research in the universities studied varied from nine to sixteen quarter hours, as indicated in Table 2. The only required element common to all four universities was some level of statistical knowledge. The use of the computer was indicated as a requirement in universities A and D only.

Table 2

Minimum Graduate Requirements in Educational Research for Doctoral Program

University	Fields of Study Expressed in Quarter Hours
A	12 - 16 hrs. Statistics Research Design Measurement Computer Application
B	12 hr. Block Research Design Statistics Methods (Action Research)
C	9 hrs. Introduction to Statistics Methods of Research
D	9 - 15 hrs. Statistics (Classical and Non-parametric) Computer Programming Survey Research

METHODS OF SYLLABI CONSTRUCTION

According to the information in Table 2, course content in each case depended upon the instructor of the course. In University B, it was indicated that course syllabi were on file in the department for the use of the instructor and for the information of other departmental personnel.

Table 3

Methods of Syllabi Construction

University	Method
A	Instructor-made (no mention of a department file)
B	Instructor-made syllabi. Have a department file
C	Instructor-made
D	Instructor-made

TEXTS, REFERENCE BOOKS, AND STYLE MANUALS

The choice of textbooks used in the four universities studied is presented in Table 4. Universities A and D emphasized textbooks on design while universities B and D introduced statistical concepts used by social scientists. The variety of texts used supported the individual freedom of instructors indicated in Table 3.

Table 4

Textbooks Used by Instructors (Doctoral Level)

University	Author	How Selected
A	Campbell & Stanley Siegel Englehardt	Instructor
B	No Specific Text	
	Masters Level: Borge & Gall (Educational Research, An Introduction) Wyatt (Statistics for the Behavioral Scientist)	Instructor
C	No Specific Text	Instructor
D	Good (Introduction to Educational Research) Heiman (Survey Design and Analysis) Tanner (Design for Educational Planning) Guilford (How to Conduct A Survey) SPSS (Stat Packs for Social Scientist)	Instructor

The chapters and subject areas emphasized from prescribed textbooks are indicated in Table 5.

Table 5

Textbooks (Chapters and Areas Emphasized)

University	Areas
A	Use plenty of books. Student chooses what he reads
B	Design chapters, Likert Scales, instrument construction, problem identification, procedures, philosophical foundations
C	Covered most chapters in text
D	Likert scale, longitudinal study

Selected books used for reference by the instructor are presented in Table 6. Each of the universities studied indicated books by Carter V. Good were primary references. Table 5 indicates specific personal preferences of the instructors. Universities A and B used Good's methods and educational research books. Table 4 on textbooks and Table 5 on instructors reference books indicate that those using books with emphasis on statistics as a text were using books on methods, designs, and The Encyclopedia of Educational Research as reference books. Those instructors using books on research methods, research design and educational research as textbooks were using books on statistics as reference books.

Table 6

Reference Books Used by Instructors

University	Texts or Authors
A	Encyclopedia of Educational Research Handbook of Research on Teaching Dictionary of Education - Good Research Methods Techniques - Good
B	Research Methods and Design in Education Educational Research - Readings
C	No Specific Reference Books
D	Guilford Curlinger Dixon and Massey - Introduction to Statistical Analysis Henry Garrett - Statistics for Psychology and Education Books by: Siegel, Morgan, Bruce

Table 7 indicates that strict adherence to a style manual was a minor consideration. Manuals by Turabian and the American Psychological Association were used most often.

SEQUENCE OF COMPONENTS IN RESEARCH PROGRAMS

Tables 8 and 9 summarize the comments of interviewees regarding the sequence of components in a research program. Table 8 is a summary of instructors' comments concerning the extent to which they felt their sequential order of course experiences was adequate. In Universities B and C instructors were dissatisfied with the arrangement of their

Table 7

Style Manuals Prescribed

University	Manual
A	No Specific Manual Use - Turabian, APA, and University of Chicago
B	Turabian Second choice - APA
C	No Specific Manual Campbell or APA
D	None Recommended APA and modified Turabian

Table 8

Reported Adequacy of Sequential Order in Research Preparation Programs

University	Course Sequential Adequacy
A	Yes, it is adequate
B	Not doing a very good job. Too much fragmentation
C	No, not adequate. Research program being reviewed by executive committee
D	Yes, it is adequate (probably most rigorous in U. S.)

Table 9

Suggestions Related to Preferable Sequences of Components for Adequate Research Programs

University	Suggested Sequences of Components for Adequate Research Programs
A	Introduction to Statistics Non-parametric Statistics Research Designs and Methods
B	Statistics Research Methods (Basic theory - then theory to practice)
C	Measurement and Evaluation Introduction to Statistics (2 sem.) Research (Intro. to) Experimental Design
D	Introduction to Statistics Advanced Statistics Research Design Computer Applications

research courses. In University C it was indicated that courses and experiences provided for students were currently under study and revision.

Research instructors in the schools studied suggested preferable sequences for teaching research courses. The results are displayed in Table 9.

TYPES OF INSTRUCTION AND INSTRUCTIONAL TECHNIQUES EMPLOYED

Table 10 indicates the types of instruction provided doctoral students in the universities studied. Only University D provided for planned simulation and interaction. Most university research classes were of the lecture type. No mention was made of field experiences except in Universities B and D.

Table 10

Percentage of Time Devoted to Types of Instruction

University	Lecture	Demonstration	Seminar
A	No specific percent. If no discussion, it would be a lecture, but there is discussion all the time		
B	50 percent	50 percent	0
C	No specific percent. All informal lectures		
D	50 percent	25 percent (Simulation)	25 percent (Interaction)

All of the instructors interviewed emphasized that they used instructional media. Instructors in Universities B and D gave most emphasis to media. However, the medium most often mentioned was the overhead projector.

Table 11

Use of Instructional Media

University	Filmstrip	Movies	Transparencies
A	None	None recommended Use of all types of media	Yes
B	None	Yes, but none good enough to recommend	Yes Instructor-made
C	None	None	Yes Instructor-made
D	Yes	None	Yes ("homemade")

All of the universities studied had computer centers that were made available to both faculty and students for the purpose of processing research data. According to Table 12, the computer in University C was available to students in advanced statistics only. In University A, one research instructor indicated that he, personally, did not make use of the computer.

Table 12

Computer Centers Available to Faculty and Students for Processing Research Data

University	Computer Center Available	Instructors Use Computer For Their Own Research
A	Yes	Yes (with one No)
B	Yes	Yes
C	Yes (Advanced Statistics Class Only)	Yes
D	Yes	Yes

In University A (Table 13), all students used computers to analyze research data; especially at the dissertation level. In University B, approximately half the advanced graduate students used a computer. It was reported in University C that all students in statistics classes used the computer, and in University D all students in the research sequence used it.

Table 13

Percentage of Doctoral Students in Educational Research Who Use Computers to Process Data

University	Percent
A	100 at Dissertation Level
B	50
C	100 of Stat Class
D	100 in research sequence

The level of understanding of computer analysis of research data is indicated in Table 14. Only in University A did faculty members do their own computer programming. In all cases, students punched cards and submitted them to the computer center for analysis.

Table 14

Extent of Computer Programming by Faculty and Students (Research)

University	Extent of Use By Faculty (Research)	Extent of Use By Students
A	2 professors in research (programming and processing)	Students at dissertation level use computer. (Punch data, submit to computer center)
B	No formal programming. Assistance from computer center personnel	Have APL Terminal in research department. No programming. Punch data cards, use "Batch" system
C	Give data to computer analyst. No programming	Students do not do computer programming. Punch data cards, submit to computer center
D	No programming. Use "canned" programs and SPSS Programs	Students do not do programming. Punch cards, submit data to computer center

Table 15 indicates that regular instruction in the use of computers was an integral part of research course experiences in Universities A and D. In Universities B and C, courses and/or seminars in computer use were available to faculty and students.

Table 15

Provisions for Teaching Faculty and Students to Use Computers

University	Method
A	Regular 3-hour course
B	Series of Seminars
C	Courses are made available
D	One course per quarter

In Table 16, instructors rated the level of cooperation between the computer center and the Department of Education from "satisfactory" in University A to "excellent" in University D.

Table 16

Levels of Cooperation Between Departments of Education and Computer Centers

University	Level of Cooperation
A	Satisfactory
B	Good
C	Very Good
D	Excellent

REQUIRED STUDENT RESEARCH ASSIGNMENTS AND STUDENT EVALUATION

All universities in the study required dissertations of all doctoral candidates (Table 17). Other major student research assignments are presented in this table. Universities A and D required survey and community assignments. More library work was stressed by Universities A and B.

Table 17

Required Student Research Assignments

University	Type of Assignment	Required to Use a Design
A	Study of person, community, or institution	Yes
B	Major proposal writing and library work	Yes Custom tailored to problem, man and occasion
C	No specific major assignment	
D	Prospectus and beyond. Define problems, determine type data needed, simulate data, research data	No particular design. Use ANOVA T-test Co-variance
	Survey Research Techniques	

Table 18 indicates a certain consistency of evaluation techniques in that such techniques were primarily subjective on the part of the instructor. Only University D made a fairly formalized attempt at objective evaluation, through a series of carefully defined and prescribed requirements and experiences.

Table 18

Bases of Student Evaluation

University	Method
A	Production, achievement, and competence in area
B	Doctoral - no exams - Active writing and participation
C	Standard tests and instructor-prepared tests
D	Final tests (teacher-made) 13 requirements must be completed. Must work with computer

OFFICES OF INSTITUTIONAL RESEARCH; THEIR RELATIONSHIP TO GRADUATE RESEARCH PROGRAMS

Universities A and D had well-developed, functioning offices of institutional research (OIR). University B had previously had such an office, but had discontinued it in favor of a faculty research committee. The office of institutional research in University C was too recently started to be evaluated as a factor in this study.

According to the data in Table 20, there seemed to be little or no functioning relationship between offices of institutional research and research programs for advanced graduate students. In University D, OIR personnel served as resource persons in the graduate educational research program.

Table 19

Universities Having Offices of Institutional Research

University	Have OIR
A	Yes
B	Discontinued. Faculty Research Committee and Graduate Dean now perform this function
C	Yes, just started
D	Yes, have two -- one for system level and one for university

Table 20

Relationships Between Faculty and Students and Offices of Institutional Research

University	Relationship - Past	Relationship - Present
A	Little contact or interaction	One student working with OIR on survey of graduate grading system
B	Do not have OIR	No OIR at this time
C	Just started	Just started
D	Little contact. OIR personnel sometimes lecture or conduct seminars for students	Provide seminars

Only in University D (Table 21) was there any working interaction between the educational research program and the OIR. One advanced graduate student was serving an internship in the OIR.

Table 21

Relationships of Instructors, Students, and Offices of Institutional Research

University	Education Instructors Work in OIR	OIR Personnel Instruct in Educ. Research Courses	Students Work in OIR	Students Are Paid for Work in OIR
A	No	No	No	Not Applicable
B	No	Not now. They have in past	No	Not Applicable
C	No	No	No	Not Applicable
D	No	No	Yes (Grad. Intern)	Yes

THE RESEARCH CLIMATES IN FOUR UNIVERSITIES

Table 22 shows that all universities in the study encouraged research and attempted to support it within the limits set by budget and available personnel.

Table 23 indicates that instructors in Universities A and D engaged in extensive off-campus consulting and field work of a research nature. In Universities B and C, this type of on-going professional involvement of instructional staff appeared to be severely limited.

Table 22

Attitudes of University Communities Toward Research

University	Attitude To Research
A	Encourages research
B	Supports and encourages research (budget slim)
C	Research encouraged. At present, research program being reviewed by executive committee. New personnel being sought for Office of Institutional Research
D	Professional community very interested in research

Table 23

Instructors Engaged in Off-Campus Consulting or Field Work of a Research Nature

University	Off-Campus Consulting and Field Research
A	Yes
B	Yes, for local school system
C	Occasionally. Evaluation of programs
D	Yes - Private Corporations Government Programs Surveys and Schools With University

Universities A and B (Table 24) required competence in or production of research for promotion of faculty members teaching research courses. In University D, field research and writing were required for promotion.

Table 24

Bases for Tenure and Promotion of Research Faculty

University	Bases
A	Expertise in teaching and research ability (some opinion that producing research aided in promotion)
B	Excellence in teaching, excellence in research
C	No specific requirements
D	Teaching, field work, writing, and other factors

SUMMARY

The preceding portions of this chapter reported what was being done to develop the research competencies of educational practitioners and researchers in the universities studied. Most programs were for training school administrators. The research programs compared in this study were established in the form of courses of study leading to graduate degrees. Traditionally, graduate programs developed for advanced training in the sciences and humanities used this approach. The tradition included a primary emphasis upon training research workers, and this emphasis was reflected nominally in the program for school administrators.

The diversity of the methods, materials, and personnel associated with different programs indicated a tendency for academic respectability to be conceded to new programs in proportion to their conformity to traditional academic practices.

It has been suggested elsewhere in this paper that there was a need to develop a program which would have flexibility sufficient to provide for a custom-made program for the researcher. The apex of support for such a program depended upon whether the nature of the researcher was basic (concerned with the discovery of new knowledge) or whether the research was applied research for the practitioner who used knowledge resulting from the basic research of others.

Campbell and others suggest three basic roles: the social scientist, the developer, and the practitioner. Since administration is an applied rather than a basic discipline, the social scientist performs for it a role analogous to that of the physical scientist to the engineer; he generates knowledge. The developer, generally in the position of professor of administration, is sufficiently familiar with the various social sciences that he may select useful knowledge from a variety of sources and transmit it to the practicing administrator.[1]

---

[1]John H. M. Andrews, "Differentiated Research Training for Students of Administration," Educational Research: New Perspectives, eds. Jack A. Culbertson and Stephen P. Hencley (Danville, Illinois: The Interstate Printers and Publishers, 1963), pp. 355-65.

Chapter 4

PROPOSED PROGRAMS FOR THE EDUCATIONAL PRACTITIONER AND THE EDUCATIONAL RESEARCHER

This chapter is devoted to an analysis of the views of experienced researchers concerning aspects of what might be considered a desirable or "ideal" research preparation program. These data and opinions were secured by direct recorded interviews and recorded telephone conversations with the professional research personnel or are quoted or paraphrased from their writings.

IDEAS RELEVANT TO THE DEVELOPMENT OF PREPARATION PROGRAMS IN EDUCATIONAL RESEARCH

In an address delivered to a meeting of the Southern Regional Council on Educational Administration, Max G. Abbott said:

> I conceive of the role of professor of educational administration as that of providing a link between the scientist and the practitioner, between those who discover and create knowledge and those who apply that knowledge to enhance practice. Thus, the role of the professor is extremely complex; so complex that it is doubtful that any one individual can perform adequately all of the functions that must be performed. In this connection, I want to make it clear that although what I have to say regarding research preparation may apply with more force to the professor who is research oriented than to the one who is practice oriented, I intend it to apply to all. Any differences that occur will relate more to depth of specific research preparation required than to the type of such preparation.[1]

---

[1]Max G. Abbott, "Research Preparation of the Professor of Educational Administration" (Atlanta: Speech delivered to the Southern Regional Council on Educational Administration, November, 1966), p. 2. (Mimeographed.)

Abbott further amplified his concept of the depth of specific research preparation in these terms:

> Adequate research design involves considerably more than an understanding of methods of statistical analysis or of historical documentation. Obviously, the researcher must acquire an understanding of theoretical formulations and conceptual schemes that enable him to organize his knowledge and to identify unexplored areas. In brief, he needs to be able to determine which questions are worth investigating.
>
> Having determined which questions are worth investigating, however, the researcher must be able to frame his questions in language that will lead to significant results. He must recognize that while language provides the avenue to the solution of scientific problems, language also serves as an obstacle to problem solution. Stereotyped language leads to stereotyped questions which give rise to stereotyped answers.
>
> . . . .
>
> It would be a mistake to assume that we are dealing here merely with a problem of semantics. It is true that students need to develop an understanding of the relation of language to events. They need to recognize that when we achieve a breakthrough in language, when new and viable concepts are developed, we are enabled to achieve a breakthrough in research. For example, we have reached a virtual impasse in our research on morale. If we are to move ahead in this general area, we will probably need to develop one or more precise concepts that will replace the concept of morale, which is now quite vague and encompassing.[2]

In 1967 at a conference on Educational Research, under the sponsorship of Phi Delta Kappa, Frank Baker said:

> I don't think that education has in fact a knowledge base of its own. Our knowledge base comes from every field except education and all those roads and all those areas lead to the mathematics department.[3]

---

[2]Ibid., pp. 2-3.

[3]Julian Stanley (ed.), Improving Experimental Design and Statistical Analysis, citing Frank Baker in a discussion on Dr. Stanley's paper, "On Improving Certain Aspects of Educational Experimentation," at the Seventh Annual Phi Delta Kappa Symposium on Educational Research (Chicago: Rand McNally and Company, 1967), pp. 45-56.

In an interview with Madison Byar, the investigator asked, "What questions in education are worth investigating?" Byar replied:

> . . . what is worth investigating is a reflection of the investigator's philosophy. What one will, or can or 'should' investigate will be a reflection of his work style; a function of beliefs. No one needs any research preparation until he has a problem; which, incidentally, will always be evidence of his faith, and, to some degree, of his most probably effective work style.[4]

Byar discussed methodology in these terms:

> Therein the problem is the question of how to solve it? That is the only time one needs methodology. How can I solve it? One needs familiarity with a lot of techniques, but doesn't need any of them until he has a problem. The 'how' of the problem necessitates the entry of research method. One has to know how to design a solution and analyze the problem--methodology comes out of the analysis.[5]

George E. P. Box, a recognized authority on statistics, presented a paper on data analysis to the Phi Delta Kappa Symposium on Educational Research. He said:

> One of the interesting things about the selection of a design is that the experimenter decides which variables, what range, how many levels, etcetera. Clearly the results of an experiment depend far more on these things than on anything else, even the data. When we design experiments, we interject our prior beliefs into the design; therefore, it's a very personal thing.[6]

---

[4]Statement by T. Madison Byar, personal interview, November 30, 1971. (Tape recorded.)

[5]Ibid.

[6]Julian Stanley (ed.), Improving Experimental Design and Statistical Analysis, citing George E. P. Box in a discussion on his paper, "Bayesian Approaches to Some Bothersome Problems in Data Analysis," at the Seventh Annual Phi Delta Kappa Symposium on Educational Research (Chicago: Rand McNally and Company, 1967), pp. 85-86.

Desmond Cook took the position that doctoral students can not all be trained to be educational researchers:

> I am not sure that you can train a person to be an educational researcher. It is possible to train a student to teach educational methodology, and anyone who goes through a doctoral program is a researcher. Anyone who wants to find the answer to a problem does research. Research deals with the question of how a person makes a judgment. The experiences a doctoral student should have are that he should do research, present it to his colleagues to be criticized, and be open to public inspection. School administrators need to understand PERT (Program Evaluation Review Technique). It is a tool, used to help manage research and development problems. As to statistics in preparation programs, the need for it depends upon what the student wants to do; and, he may not know. If his interest is in the quantitative approach, the more statistics he has had the better off he will be.
>
> Today some schools are administered, some are managed. Management is more sophisticated, if studied statistically, relative to cost benefits; a kind of accountability. Administration is more of a maintenance operation; the administering of a program someone else established.
>
> A student who would use PERT will need a knowledge of measurement but not necessarily of research design.[7]

When asked if graduate students of education should have knowledge of and use computers, Cook said, "Yes, he should run through the whole business of the program."[8]

When asked about the reporting of articles in research journals, Cook replied that, "the articles reported in the Journal of Educational Research constitute a kind of sterile approach."[9]

Arthur DeRosier discussed research preparation programs and historical research. He said:

---

[7]Statement by Desmond Cook, telephone interview, January 27, 1972. (Tape recorded.)

[8]Ibid.

[9]Ibid.

Historical research, to me, is the most exciting type of research. One gathers data about events that happened years ago, and from this small amount of data projects what actually happened. One cannot escape the age in which he lives, the ideals he has, and the contemporary scene. When one gathers data it may all be correct, but it tends toward his side of the particular issue.

One difficult task of the student of historical research is synthesizing. One can synthesize something only when he knows more than what he is trying to synthesize. One should write out of a knowledge of the material.

I don't think a doctoral student can perform historical research without bringing the competencies (to do research) with him. To be a good researcher one must do it (research) and do it again and again.

Flexibility is needed at all levels of education. The regular class schedule is geared to passing on the heritage. It is not geared to student participation. The doctoral student planning to become a school administrator should be exposed to a number of courses in research. Whether historical research or educational research, the same kind of informality and group interplay must figure into the training program.[10]

Robert M. Gagné discussed one of the most difficult problems facing the educational researcher--"that is, of accomplishing the goal of getting people to know how to define problems."[11]

Gagné further asked the question: "How does one provide the understanding of experimental design and statistics and so on that is needed to make people avoid doing incorrect kinds of things?"[12]

---

[10]Statement by Arthur DeRosier, Jr., personal interview, February 23, 1972. (Tape recorded.)

[11]Julian Stanley (ed.), Improving Experimental Design and Statistical Analysis, citing Robert M. Gagné in a discussion on Dr. Stanley's paper, "On Improving Certain Aspects of Educational Experimentation," at the Seventh Annual Phi Delta Kappa Symposium on Educational Research (Chicago: Rand McNally and Company, 1967), p. 57.

[12]Ibid., p. 47.

Alfred Garvin discussed doctoral level preparation programs in educational research and said, "It would take three years of hard study to learn the tools of research. It would take all one's life to learn to be useful."[13]

Carter V. Good, a nationally known authority on educational research, said:

> Good research reveals the humanism in people. The order of courses is difficult to determine; though, make certain in a three-year doctoral program, to include Introduction to Research Methods, plus specialization.[14]

Good included in his remarks that the student of educational research should have knowledge of the chief research approaches: introduction to major tools and methods of problem solving; knowledge of historical methods; introduction to survey techniques (picture of present methods); introduction to computer techniques (ways of analyzing data). Researchers have learned a great deal from other disciplines concerned with surveys (anthropology and sociology).[15]

Good further stated:

> The student of educational research should study statistical techniques, basic techniques, and introduction to logic, theory, and inquiry. Analysis of data should precede experimental design. An introduction to historical methods and a background of educational history is important. I have been interested in bibliographical work, historical summarizing and synthesizing research studies over a number of years--the historical approach.[16]

---

[13]Statement by Alfred Garvin, personal interview, November 30, 1971. (Tape recorded.)

[14]Statement by Carter V. Good, personal interview, November 30, 1971. (Tape recorded.)

[15]Ibid.

[16]Ibid.

Egon Guba discussed writing of proposals and said:

> Almost all studies have two potential audiences--practitioners and professional peers. What will the research mean to other psychologists and what will it mean to school teachers in the classroom?[17]

John Hoyle said:

> There should be a sequence of educational research courses for preparation of students. We should develop a little pride in education; have as much strength as other social scientists. The sequential order should be put in a block form: design, statistics, and research all together. The student should develop a design. There should be a sequence. We are not doing a very good job of sequencing. There is too much fragmentation. Where are you? If you don't know, any road will take you.[18]

John Lovell suggested experiences for training educational researchers should include statistics and research design. He (doctoral student) should be a producer of research.[19]

Ralph Purdy discussed methodology and research training in these terms, "I am not interested in methodology until I have a problem. There is the problem--how to solve it--this is the only time you need methodology."[20]

Purdy extended his philosophy and opinions on educational research in these terms:

---

[17]Egon G. Guba, "The Writing of Proposals." (Mimeographed.)

[18]Statement by John Hoyle, personal interview, December 1, 1971. (Tape recorded.)

[19]Statement by John Lovell, personal interview, January 13, 1972. (Tape recorded.)

[20]Statement by Ralph Purdy, personal interview, December 1, 1971. (Tape recorded.)

For what kind of problem faces the administrator for which he needs data--this gives direction to the design. There is the man, the problem, and the ultimate result. Design has to be 'custom-tailored' to the man, the problem and the occasion. I am not ready to begin the design of my house until I know the lot on which it will be placed. The design is supposed to produce data and is developed to achieve objectivity of the data of the problem. I have never been able to design except in relation to the problem.

The type of course offering should contain action research and teach one to be an interpreter and user of research findings, problem identification, design research techniques, procedures and philosophical framework. The prospective researcher should have a good overall understanding of research. There is a difference in involvement in proposal writing and individual study--show and tell, wit and 'git'.

Courses should be aimed toward basic beliefs and values on the pattern of personal lives. They should clarify that which is defensible and look at tools, by which the researcher implements. The basic courses are theory, philosophy of education, and theory and practice (seminar). The program in educational research is suited to the type of research he is doing, not pure research as such--action research.[21]

Julian Stanley discussed nurturing educational researchers, and said:

It seems to me that after we have produced a well-prepared research-oriented doctoral recipient we should nurture him carefully for ten or fifteen years to be sure he has every incentive to do excellent, continuing research. I have been trying for years to emphasize the necessity of doing experiments in education where the types of answers we are trying to get require experiments, or at least lend themselves very well to the experimental approach. I think we have frozen the educational researcher into a psychometric camp. Perhaps by 1975, 75 percent of all educational research studies will be experiments.[22]

Stanley discussed a sequence of courses in these terms:

We don't set up sequences or programs of course as such for training researchers. Such courses as mathematical

---

[21]Ibid. [22]Stanley, op. cit., pp. 22-23, 30.

> statistics, matrix theory, and probability theory are recommended. One of the greatest things in promoting this particular type of program at Wisconsin has, I believe, been the extreme flexibility of course choice.[23]

C. Kenneth Tanner said:

> The prospective educational researcher needs twenty-five to thirty semester hours of training to sell himself as a researcher. The sequence of courses should include Introduction to Statistics, Advanced Statistics, and Research Design.[24]

Tanner placed emphasis on developmental research--developing models. He said doctoral students should be working in doctoral research internships with professors.

A doctoral student, who wished to remain anonymous, discussed research preparation and said:

> I have a fear of research design. It would help when it is more meaningful and relevant. I have had as much statistics as I would like to have (two courses) at this time--it is too threatening; a non-personal type lecture course.[25]

The student commented on the use of the computer, and said, "If for no other reason than getting maximum efficiency, hands-on use of the computer would be helpful."[26]

THE PROGRAM: AUTHORITY TO ORGANIZE, PLAN, PROGRAM, AND OPERATE (OPPO)

"In any effective form research is power, just as clear thinking about real problems has always been a source of human

---

[23]Ibid., p. 59.

[24]Statement by C. Kenneth Tanner, personal interview, January 13, 1972. (Tape recorded.)

[25]Statement by a doctoral student, personal interview, December 1, 1971. (Tape recorded.)

[26]Ibid.

influence."[27] A training program for a doctoral student should include the significance of research in the administrative power structure; that is, the authority to permit the researcher to perform research or initiate projects. Therefore, realistic concepts as in the Theory of Tripartite Power should be included in a student's custom-made program. The line of authority to organize, plan, program, and operate (OPPO) should be studied, not in isolation, but rather in actual practice.

A peripheral analysis made by individual research professors at the institutions in this study confirmed a strict adherence to vertical lines of authority and echelons of power and control.

How, then, would one design a preparation program to encompass exposure to the different methods, experiences, personalities, emphases, and individualities discussed "in-person" by the researchers and administrators included in this study?

The following design was woven from little pieces contributed by each of the research authorities interviewed or quoted. The design was based on the premise that each person has an individual way or style of performing work or "doing the job." Work-style is then framed in a total structure of research preparation from organization to planning, programming, and operation (OPPO).

It was a major assumption of this study "that extant university research programs for the preparation of school administrators and educational research personnel reflect the philosophy, beliefs, and work-styles of those who plan and operate them." That is, the way research and research training is Organized, Planned, Programmed, and

---

[27]Wilson, Byar, and others, op. cit., p. 338.

Operated can be expressed by the acronym OPPO, which will here be used to represent systems of research and research preparation. Each type of program will yield to OPPO analysis and the activity resultant from an OPPO will necessitate an identifiable work-style. It is, therefore, imperative that a researcher's preferable work-style be determined before he undertakes advanced research preparation or field work of a research nature.

Work-style was defined as being founded in an individual's beliefs, "by those operational factors in his thinking that determine both what types of things he is most likely to do and the way or ways he is most likely to go about doing them."[28] People who successfully carry out research planning, design, and operations have "a bent"; they lean toward faith in the results of mathematical calculations and measurement, toward the philosophical ideas that prompt cultural developments, toward surveys of status quo to explain and/or rationalize the need for change or no change, toward experimental approaches as problem-solving devices, or toward the analyses of trends or opinions of apparent results. In any case, few if any individuals possess the broad variety of interests or expertise to reach top competence in very many of those areas of interest or the work-styles they require. One element, then, that is foremost in the determination of any individual's research preparation needs is an analysis of the way he works and of what, if it is possible to know, makes him work best in that way. If and when students and their advisors squarely face these questions it

---

[28]Statement by T. Madison Byar, personal interview, January 20, 1972. (Tape recorded.)

will become apparent what sorts of problems may seem researchable to the student, and those problems will determine what research techniques he will see as promising or that he will need to know.

CUSTOM-MADE RESEARCH PROGRAMS

Both an individual student's goals and the nature of his operational beliefs should be considered in planning his research preparation program. Perhaps such considerations can be undertaken only by the "planner type," the generalist who must consider all types of questions, possible research techniques, the results of the uses of various techniques and the reasons, both personal and scientific, for using them. Other individuals generally need other types or perhaps a more limited type of preparation, suited to their most likely work-style.

A preparation program should be reality oriented. That is, the student should be "on-the-job," "in-actual-practice," "doing research." A good program should contain provisions for "flexible studying," for organizing, planning, programming, and operating research related activities.

The program herein proposed would allow each doctoral student actually to work in a school system, or research framework, to expose himself to a variety of experiences (someone or some group) to learn the techniques and possibilities of research.

An internship with one working in research is probably the best medium through which to develop a student's work-style. His experience may be either positive or negative. That is, he may "fit in" or "fall out" because of the beliefs he took to the job--or he may actually change beliefs on the basis of new experiences. However, there is no better way to find out than to "farm him out."

An OPPO system of research training would be as follows:

1. Each doctoral student would be required to demonstrate recognized competency in research techniques.

2. A research competency once recognized would be developed in a real-life situation.

3. The amount of credit to be granted and the length of the research preparation should be determined by the student and his major advisor. All student's research preparation should include an Introduction to Research Methods, Statistics, and work in the Writing and Interpreting of Research Reports. After nine quarter hours of credit in these areas the rest should be a variable custom-made program for those seeking or needing advanced preparation.

4. In any case, each doctoral student should be administered a "Student Work-Style Orientation Index" (to be developed) to use with other measures in determining work-style by discovering preferences for doing things and also checking other opinions on a scale (Likert type) to see whether the student's preferences showed up in his performance.

5. The Student Work-Style Orientation Index would be available to the major advisor or coordinator of the OPPO Research Program at the University.

6. Individual public or private school systems (elementary, secondary, and collegiate) may request doctoral students in the OPPO training program to perform as researchers for their systems or institutions. The chief administrators of the systems should describe the areas or types of research they wanted performed.

The doctoral student researcher would be available at no fee to school systems to perform research functions. The contribution from the individual school systems should be actual employment status as researchers. Monetary compensation would be replaced "in-kind" by making available to the "newly-employed" faculty member or administrative rank researcher, the facilities, machinery, and above all, individual administrators or teachers and other personnel with whom the researcher would work. The doctoral researcher obtained through the university would not be considered as a student and given only "token" time by individual administrators or teachers with whom he worked. Rather the doctoral researcher would receive the same treatment given to other personnel in similar roles or responsibilities. Therefore, administrative and teacher "released time" to work and help the researcher would be considered as compensation to the doctoral researcher in the form of experiences that he would never be able to buy.

7. The request by an individual public school superintendent for a particular type of research worker would immediately be answered by referring to the Work-Style Research Orientation Index and other measures from the doctoral student files.

Prior to this, a list of available doctoral researchers and the types of work they could and would prefer to do would be provided the school systems in the OPPO program.

If, for example, the school superintendent wanted to conduct research on his reading program he would select from the list the available doctoral students whose work-style reflected a preference for and training in the area of reading.

8. It is to be understood that the doctoral student would have completed a minimum of nine quarter hours of research preparation before being included in the list for actual employment as an intern researcher. In many situations the nine hours might have been completed at the Masters' degree level. The choice of nine quarter hours is to satisfy accreditation standards, which are expressed in quarter hours of credit.

9. The completion of courses in research methods, statistics, and writing and interpreting research would be on an individual basis developed by the student and his major advisor on the initial description of his custom-made program.

10. Different amounts of credit would be allowed, depending upon the depth of research preparation desired, and might include: (1) knowledge of and ability to handle statistical procedures as a prerequisite; (2) introduction to techniques of research; (3) evaluating, interpreting, writing, reporting research findings; and (4) advanced research and/or specialization, including special research designs, advanced statistics, comprehensive field surveys, and computer applications.

As a result of the findings from the literature and the interviews, it was determined that the doctoral student majoring in school administration needed some expertise in methods of research, in writing and interpreting research, and in advanced research. A syllabus was developed to include courses and experiences to assist advanced graduate students to develop the expertise needed by the practitioner of educational research. A brief outline is presented in this chapter, including a course description, behavioral objectives, and a syllabus design. Volume II, Chapters 1, 2, and 3 give more detail concerning the courses. The program is outlined in a block of courses and experiences linked to a real research problem likely to confront the practitioner.

The program is designed to be conducted by three distinct but related groups: (1) the Peer Group, (2) the Reference Group, and (3) the Resource Group. Each group would be composed of four members. The Peer Group would consist of advanced graduate students who would serve as counselors and advisors to their peers. The Reference Group would consist of professors, school administrators, and advanced graduate students. The reference librarian would be invited or recruited as a key member of this group. The problem, the man, and the occasion would be used to determine the selection of Reference Group members. The Resource Group would be made up of professors. The group would have members with expertise in research design, statistics, psychology and sociology, and computer programming and operation. This group would be available throughout the program to help students with their problems.

The block course would be conducted by the Peer Group, the Reference Group, and the Resource Group to provide immediate help to advanced graduate students having practical problems resultant from real research experiences. Flexibility of program would be provided to allow for individual differences and personalized learning. A team approach involving an educational research professor and the other groups would be used.

A course syllabus describing research experiences would be offered advanced graduate students. A proposed syllabus follows.

RESEARCH METHODS

Textbook

A textbook would be chosen by the professor. A bibliography in the course syllabus would serve to locate reference material.

Course Description

This course deals with the application of the knowledge and methods of educational research to problems which may serve as thesis and dissertation topics in a graduate student's program. While it is recognized that students are not experts in research methodology, it is expected that students will develop sufficient understanding of the criteria used to evaluate research to be able to critically analyze research which has been completed, to appraise the potential of proposed research topics, and to develop defensible research proposals of their own.

Topics Covered

A. Internal and External Validity

B. Research Terminology

C. Research Designs

Pre-Experimental Designs:

1. A One-Shot Case Study
2. A One-Group Pretest-Posttest Design
3. A Static-Group Comparison

True Experimental Designs:

4. A Pretest-Posttest Control Group Design
5. A Solomon Four-Group Design

6. A Posttest-Only Control Group Design

Quasi-Experimental Designs:

7. A Time-Series Experiment

8. An Equivalent Time-Samples Design

9. An Equivalent Materials Design

10. A Nonequivalent Control Group Design

11. Counterbalanced Designs

12. A Separate-Sample Pretest-Posttest Design

13. A Separate-Sample Pretest-Posttest Control Group Design

14. A Multiple Time-Series Design

15. A Recurrent Institutional Cycle Design: A "Patched-Up" Design

16. Regression-Discontinuity Analysis

17. Questionnaire

18. Interview

19. Survey

20. A Q-Sort

D. Bibliography of Research References (Volume II)

Behavioral Objectives

1. Given a report of completed research, the student will be able to analyze and appraise the research using pre-established criteria as standards.

2. Given proposals for research, the student will be able to critically analyze the potential as well as the limitations of the design and will be able to offer suggestions for improving the design if warranted.

3. Given a list of terms concerning research design, the student will be able to define the terms orally or in a written narrative form. Seventy-five percent of terms correctly defined will be judged as a success.

4. Given a list of designs, the student will be able to match the design to the statistical technique appropriate to it.

5. Given a bibliography of readings concerning design, the student will do supplementary reading in the library.

Evaluation

Evaluation will be based on the behavioral objectives and will include the following:

1. Each student will present oral reports and have interactive discussion on all of the different designs (in capsule form).

2. A test will be administered (if necessary) consisting of questions about the designs.

3. A debate will be conducted where students discuss the designs.

Materials

1. "Handouts" of each research design for each student
2. Copies of terms and reference materials for each student
3. Transparencies, filmstrips, and other instructional aids (if desired)

STATISTICAL TECHNIQUES

Textbook

A textbook would be chosen by the professor. The bibliography (Volume II) would serve to locate reference material.

Course Description

This course is designed to familiarize students of educational research with the literature pertaining to statistics, and with the statistical techniques which match research design and data.

Topics Covered

1. Analysis of Variance
2. Analysis of Covariance
3. Factorial Analysis of Variance
4. "t" Test
5. Mixed Designs
6. Chi-Square (X^2)
7. Reliability
8. Orthogonal Comparisons
9. Tchebychef's Formula
10. Alpha-four
11. F-Ratio
12. Pearson-Product Moment
13. Regression Analysis
14. Introduction to Bayesian Statistics

Behavioral Objectives for Statistical Techniques

1. Given handout sheets on statistics and a bibliography containing references on statistical techniques, the student will be able to discuss the techniques orally or written in narrative form. The discussion will be evaluated on a seven-point Likert scale. The student should average four or better on the Likert scale.

2. Given problems dealing with statistical techniques, the student will be able to work them for assignment. The evaluation will be based on 100 percent; a score of 70 percent will be acceptable for minimal success.

3. Given a list of terms describing statistical techniques, the student will be able to develop acceptable definitions of the terms. This list will be checked by the instructor as an assignment exercise. Those definitions which need improvement will be checked and the list will be returned to the student for improved definitions.

4. Given a list of terms to define, the student will be able to define the terms at least 75 percent of the time.

5. Given a list of statistical techniques to match with research designs, the student will be able to match at least 75 percent of the pairs.

6. Given the same matching exercise as in 5 above, the student will be able to discuss in narrative form the reason for choosing the research design and matching research technique. Seventy-five percent accuracy will be required.

Evaluation

Evaluation will be based on the behavioral objectives and will include the following:

1. Each student will work problems on each of the topics.

2. Oral reports and interaction will be required to demonstrate understanding of each problem and statistical technique.

3. Each student will demonstrate competence in solving each problem and in discussion of each topic.

Materials

1. "Hand-out" of statistics problems

2. Copies of list of terms

3. Transparencies and other instructional media as required

COMPUTER APPLICATION

Textbook

A textbook would be chosen by the professor. STAT PACS for Social Scientists and Introduction to FORTRAN would also be required. The 1130 Computer Monitor Manual would be required as reference material. The bibliography (Volume II) would serve to locate reference material.

Course Description

Computer Application is a course designed to give the student an opportunity to understand computer usage and programming. Correlatives of this course are an understanding of the Key-Punch Machine, the arrangement of data on the key-punch card, and a complete functional understanding of the research designs and methods used for data analysis.

Topics Covered

1. Preparation of Data
2. The Control Cards
3. Correlation Analysis
4. Variance Analysis
5. Non-Parametric Statistics
6. Regression Analysis
7. Computer Operation
8. Error Analysis ("de-bugging")

Behavioral Objectives for Computer Application

1. Given handout sheets concerning computer application to statistical procedures, the student will be able to key-punch the program for the subroutines and the data cards required. Success will be measured by the correct answer on the computer print-out.

2. Given correct program and data cards, the student will be able to put the computer cards in the 1130 computer and receive a print-out of the correct answer.

3. Given handout sheets concerning research design and computer application to statistical procedures, the student will be able to match the statistical procedure to the research design. A score of 75 percent or above will be evaluated as success.

4. Given a list of terms on computers, the student will be able to discuss them orally. Those definitions not acceptable will be revised by the student.

5. Given a matching exercise involving computerized statistical techniques and research designs, the student will be able to discuss

orally, or write in narrative form, the reasons for his choices. A score of 75 percent or above will be judged satisfactory.

6. Given a research case study, the student will be able to decide on a research model, write a computer program in FORTRAN language to analyze the data, key-punch the program cards, the control cards, the data cards, and complete a successful print-out of the findings. A final report consisting of analysis of findings of the research study completes this assignment.

Evaluation

Evaluation will be based on the behavioral objectives. Each student is required to have "hands on" operation of the computer.

Materials

1. The handout Computer Application Manual and Instruction (CAMI)
2. Time on the 1130 Computer
3. Cards and FORTRAN coding sheets
4. Key-punch, Disc Packs, Sorter
5. Library of Subroutines
6. Control cards
7. Reference materials

WRITING AND INTERPRETING RESEARCH

Textbook

A textbook would be chosen by the professor. A bibliography (Volume II) would serve to locate reference material.

Course Description

This course was developed for the purpose of giving students an opportunity to develop a defensible prospectus for a research topic. The course is designed to review, analyze, and interpret the most applicable research findings in several major areas of interest to educational administrators. It is not aimed at a single research topic in depth, but at presentation in breadth and interpretation of research findings relating to the role of the educational administrator.

Course Outline

A. Overview of educational research
B. Discussion of Research Analysis Rating Scale
C. Discussion of the writing of a proposal
D. Schedule of students' analyses of written documents for their research value
E. Schedule of student's oral presentation of his prospectus and his defense of it before his peers and professors

Topics Covered

Research studies dealing with an appropriate selection from:

1. Community support for education; election involving school issues
2. Decisive factors in potential strike situations
3. Evaluation of school-community relations
4. Approaches to school-community relations
5. Public opinions of the schools, colleges, and universities
6. Trends in teachers' strike activity

7. Teacher turnover; why teachers and professors leave their jobs
8. Public expectations of boards of education
9. Teacher interview and selection techniques
10. Decision-making and budgets
11. Decision-making and systems analysis
12. Power structure relation to school boards
13. Decision-making by groups and individuals
14. Communicating with the public
15. Educational planning
16. Accountability

Objectives

1. Identify a problem which can be developed as a dissertation topic.
2. Define the problem in such a way that is adequately researchable.
3. Become familiar enough with analytic procedures to continue development of the problem (Logic and statistics).
4. Become familiar with aspects of measurement which are appropriate to the problem.
5. Study proposal writing styles and forms with the intent of adopting aspects of them for the proposal.
6. Become aware of university resources available for help (Library--ERIC, Education Index, Dissertation Abstracts).
7. Learn the departmental requirements for the submission of the dissertation proposal.
8. Complete a dissertation proposal.

Activities

1. Pre-testing over basic knowledge of statistics, measurement, design and administrative research

2. Learning research methodology, design, and implementation through lectures and readings

3. Writing and reporting various stages of the proposal

4. Discussing strengths and weaknesses of proposal problems

5. Literature searches to demonstrate familiarity with the library

6. Conferences with resource group

7. Reporting to group the progress made with the committee toward acceptance of the proposal

8. Discussion of implementation tactics of research findings

9. Student presentation and discussion of selected topics

10. Definition of other needed research in educational administration

11. Writing research findings in laymen's terms

12. Each student will be required to submit five (5) research reports. That is, each written research report will be reproduced and distributed for discussion.

<u>Steps Needed for the Interpretation of Research</u>

1. Select a topic that involves an educational problem.

2. Obtain the research from the suggested references below and other references (Volume II).

3. Write the report according to the suggested format.

<u>References</u>

1. Journals

2. Dissertation Abstracts

3. Educational Administration Abstracts

4. ERIC and ERIC Abstracts

5. <u>Review of Educational Research</u> (Journal)

6. <u>American Educational Research Journal</u>

7. Federal Project reports

8. Books

SUGGESTED FORMAT FOR REPORTING RESEARCH
TO PRACTITIONERS

Title of Topic

Introduction

One or more paragraphs introduce the nature of the research that is included in the analysis and interpretation.

Questions

What questions did the research address?

Review of Studies

Include the following points about each study:

a. Where the study was conducted
b. Who conducted the study
c. Purpose of the study (should relate to one or more of the questions above)
d. Sample size (if appropriate) and regions or state investigated
e. Who participated in the sample (administrators, teachers, professors, etc.)

Conclusions and Implications

Note that your job is to keep the research reporting in simple terms. Try to write the report for persons who have completed from twelve to fourteen years of formal education.[29]

---

[29] C. Kenneth Tanner, "An Outline for Analysis and Interpretation of Research for Educational Administrators" (Knoxville, Tennessee: University of Tennessee, 1971), pp. 2-3. (Mimeographed.)

SUGGESTED PROGRAMS FOR EDUCATIONAL RESEARCHERS

The program for the educational researcher will be more comprehensive and will be aimed at developing expertise in basic research as well as applied research. The student pursuing this program might fill the roles of change agent, county agent, educational development specialist, retriever converter, field tester, and/or quality control man. The suggested outline follows. The program for educational researchers will also be presented in Volume II, Chapter 4, for comparison with the program for the educational practitioners.

Program for the DOCTOR OF EDUCATION Degree (Major in Educational Research)

Area		Qtr. Hrs.
Statistics	Intro. to Statistical Methods	3
	Intermed. Statistical Methods	3
	Nonparametric Statistics	3
	Multivariate Statistics	3
	Intro. to Bayesian Statistics	3
Evaluation	Measurement and Evaluation	3
	Adv. Measurement and Evaluation	3
	Surveys and Rating Scales	3
	Theory of Measurement	3
Research Methods	Research Methods and Techniques	3
	Research Design	3
	Adv. Research Design	3
	Survey Research	3
	Intro. to Educational Planning	3
	Indiv. Study in Ed. Research	6
	Writing and Interpreting Research	3
	Computer Programming - FORTRAN	3
	Federal Programs & Grants	3
	Adv. Research & Analysis	3
Other Courses in:	Social Foundations of Education	3
	Curriculum or Media	3

Area	Qtr. Hrs.
Two minor fields and/or cognate work in Psychology, Sociology, Mathematics, etc.	48
Dissertation	24
	138

Chapter 5

SUMMARY OF FINDINGS, CONCLUSIONS, AND IMPLICATIONS

FINDINGS

It was the purpose of this study to determine whether university preparation programs bore out the findings of critics, and to synthesize opinions and facts regarding research into a proposal for individual custom-made preparation programs. The procedure involved actual visits to universities to study their programs; a review of literature; and actual live (recorded) interviews with certain outstanding educational research personnel, professors, and administrators. These three approaches: (1) reading of what has been done in educational research programs; (2) actual visits to university programs; and (3) talking "in-person" through recorded interviews with some of the leaders in education, were considered adequate.

Each of the research authorities described a plan or procedure for research preparation programs. These authorities did not or would not attempt to prescribe a sequence of courses for an adequate program. Each held to the traditional research methods, statistics, and measurement courses offered by all universities having doctoral programs. Yet, in each instance, they strongly recommended that a proper research preparation program could and should be developed wherein the student actually worked as an intern to someone engaged in educational research.

It appeared paradoxical that these university professors taught courses for preparation in educational research as listed in their university catalogs while privately advocating actual research involvement as the best type of training. In addition, the question of what type of student best qualified to be a researcher was argumentative. Discussion about whether the student should be competent in mathematics prior to educational research preparation was included in almost all interviews. The authorities knew what was needed in a student's research preparation program, but they could not identify a method or vehicle to use in designing a preparation program to allow for differences in education, background, style, and approach of doctoral students.

CONCLUSIONS

A major conclusion of this study is related to Byar's statement concerning the "bedevilment theory"; that one cannot get the man out of his research. All those involved in this study revealed a personal approach to educational research and research training. That is, any course is like the professor who teaches it.

Another conclusion was that research preparation for the practitioner must be "custom-made" for competency in consuming research and developing decision-making skills. It is around "decision research" that the practitioners' research preparation program should be built.

IMPLICATIONS

The dilemma of research preparation programs, then, constituted a major implication of this study. That is, each student has his own unique "work-style." His program should be "custom-made." As Desmond

Cook said, "What does he (the student) want to do? He may not know. If he wants the quantitative approach, the more he has the better off he will be."[1]

A "custom-made" educational research preparation program should be based on a student's "work-style." That may not be simple but it is an individual way to prepare a student for research. That thesis will be considered in Volume II of this study.

---

[1]Statement by Desmond Cook, telephone interview, January 27, 1972. (Tape recorded.)

BIBLIOGRAPHY

BIBLIOGRAPHY

Volume I

A. BOOKS

Buswell, Guy T. and T. R. McConnell. Training for Educational Research. Cooperative Research Project No. 51074, University of California, Berkeley, 1966.

Clark, David L. and John E. Hopkins (eds.). "An Analysis of Formal Programs for the Training of Educational Research and Dissemination Personnel," A Study of Roles for Researchers in Education. Bloomington, Indiana: School of Education, Indiana University, 1969.

Culbertson, Jack A. and Stephen P. Hencley (eds.). "Differentiated Research Training for Students of Administration," Educational Research: New Perspectives. Danville, Illinois: Interstate Printers and Publishers, 1963.

________. "Educational Research; a National Perspective," Educational Research: New Perspectives. Danville, Illinois: Interstate Printers and Publishers, 1963.

Fattu, Nicholas A. "A Survey of Educational Research at Selected Universities," First Annual Phi Delta Kappa Symposium on Educational Research. Bloomington, Indiana: Phi Delta Kappa, 1960.

Johnson, Palmer O. "Introductory Remarks at Opening of the Symposium on Educational Research," First Annual Phi Delta Kappa Symposium on Educational Research. Bloomington, Indiana: Phi Delta Kappa, 1960.

Krathwohl, David R. "Current Formal Patterns of Educating Empirically Oriented Researchers and Methodologists," The Training and Nurturing of Educational Researchers. Bloomington, Indiana: Phi Delta Kappa, 1965.

Monroe, Paul. Encyclopedia of Educational Research. New York: The Macmillan Company, 1969.

Sax, Gilbert. Empirical Foundations of Educational Research. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.

Sieber, Sam D. Analysis of USOE Research Training Programs, Bureau of Applied Social Research. New York: Columbia University, 1968.

Stanley, Julian. Improving Experimental Design and Statistical Analysis, Seventh Annual Phi Delta Kappa Symposium on Educational Research. Chicago: Rand McNally and Company, 1967.

Travers, Robert M. W. An Introduction to Educational Research. New York: The Macmillan Company, 1958.

Wandt, Edwin. A Cross-Section of Educational Research. New York: David McKay Company, 1965.

Wilson, L. Craig, T. Madison Byar, Arthur S. Shapiro, and Shirley H. Schell. Sociology of Supervision, An Approach to Comprehensive Planning in Education. Boston: Allyn and Bacon, Inc., 1969.

B. PERIODICALS

American Psychological Association. "Report of the Seminar on Education for Research in Psychology," American Psychologist, XIV (April, 1958), 167-69.

Coladarci, Arthur P. "Toward More Rigorous Educational Research," Harvard Education Review, XXX (1960), 3-11.

Stanley, Julian. "Preparing Educational Research Specialists for School Systems: A Proposal," Phi Delta Kappan, XLVIII (December, 1966), 110-13.

C. UNPUBLISHED WORKS

Abbott, Max G. "Research Preparation of the Professor of Educational Administration." Speech delivered to the Southern Regional Council on Educational Administration, Atlanta, November, 1966. (Mimeographed.)

Altschuld, James William. "A Study of an Experimental Training Program in Educational Research and Development: The Measurement and Analysis of Factors Predictive of Graduate Student Success." Unpublished Ph.D. dissertation, The Ohio State University, 1970.

Byar, T. Madison. Personal interview. November 30, 1971. (Tape recorded.)

Byar, T. Madison. Personal interview. January 20, 1972. (Tape recorded.)

Cook, Desmond. Telephone conversation. January 27, 1972. (Tape recorded.)

DeRosier, Arthur, Jr. Personal interview. February 23, 1972. (Tape recorded.)

Fleury, Bernard J., Jr. "A Study of Factors Relevant to the Development of Applied Educational Research Training Programs." Unpublished Ed.D. dissertation, University of Massachusetts, 1968.

Garvin, Alfred. Personal interview. November 30, 1971. (Tape recorded.)

Good, Carter V. Personal interview. November 30, 1971. (Tape recorded.)

Graff, Orin B. Telephone conversation. January 27, 1972. (Tape recorded.)

Guba, Egon G. "The Writing of Proposals." Bloomington, Indiana: Indiana University, n.d. (Mimeographed.)

Houston, Clifford James. "Trends in Research Tool Requirements for the Doctorate in Education." Unpublished Ph.D. dissertation, Colorado State College, 1969.

Hoyle, John. Personal interview. December 1, 1971. (Tape recorded.)

Lovell, John. Personal interview. January 13, 1972. (Tape recorded.)

Millikan, Nancy Horne. "Processes of Socialization in Educational Research: Delineation of Problems Facing the Development of Researchers by Graduate Institutions of Education." Unpublished Ph.D. dissertation, Columbia University, 1969.

Puffer, Richard Judson. "The Educational Research Involvement and Capabilities of Institutions for Teacher Education." Unpublished Ph.D. dissertation, Northwestern University, 1967.

Purdy, Ralph. Personal interview. December 1, 1971. (Tape recorded.)

Sakumura, Joseph Suyeo. "An Analysis of an Experimental Research and Development Program for Talented Undergraduate Education Students." Unpublished Ph.D. dissertation, The Ohio State University, 1969.

Sproull, Natalie Loraine. "The Relationship of Research Preparation to Duration, Attrition and Research Communication of Doctoral Candidates in Education." Unpublished Ph.D. dissertation, Michigan State University, 1969.

Tanner, C. Kenneth. "An Outline for Analysis and Interpretation of Research for Educational Administrators." Knoxville, Tennessee: University of Tennessee, 1971. (Mimeographed.)

________. Personal interview. January 13, 1972. (Tape recorded.)

APPENDIXES

Volume I

APPENDIX A

FACULTY INTERVIEW RESPONSE DATA FORM

1. Name______________________________ Rank____________________

   Address____________________________ Department________________

   Telephone___________________________ Yes in Dept. at this Univ.____________

   Total years of Experience_______________

2. How many instructors regularly teach educational research courses?

3. What are the minimum courses required in educational research beginning at the Master's level and continuing through the doctoral level?

   Total quarter hours?

4. What textbooks do the instructors use?

   How selected? Instructor Other

5. What style manuals are prescribed?

6. Name three (3) favorite reference books used by the instructor in addition to the text which he relies heavily to supplement the textbook?

7. What filmstrips do the instructor use (if any)?

8. What movies do the instructor use (if any)?

9. To what extend do the instructors use transparencies? Are they commercially made or instructor made?

10. To what extent are the research courses' lecture demonstration as opposed to seminar and student discussion?

11. What kinds of major research assignments do students have in each research course?

    Is each student required to use, develop, and apply a particular design or designs as a part of his course experience?

12. On what basis does the instructor evaluate students in research courses?

13. What is the attitude of your university toward professional research?

14. What are the bases of tenure and promotion in this university?

15. How are course syllabi constructed?

16. Indicate which chapters of the textbooks the instructors use in research classes?

17. What is your opinion of the adequacy of the sequential order of research courses, topics, and experiences throughout the entire educational research program?

18. Is there a sequence in the components of an adequate program of educational research? If so, what?

19. Is the computer center available to faculty and graduate students for processing research data?

20. Do you use the computer personally?

21. Approximately what proportion of students in your educational research classes use the computer to process data?

22. Do faculty and students do their own computer programming? To what extent?

23. What provision is made to teach faculty and students computer programming?

24. What would you say about the level of cooperation between the department of education and the computer department?

25. Does the university have an office of institutional research?

26. What is the relationship between faculty and students and the office of institutional research?

27. Provide a couple of examples within the last three or four months in regard to question 26 above.

28. Does any instructor of educational research work part-time in the office of institutional research?

29. Does any researcher in the OIR instruct part-time in educational research?

30. Does any graduate student in the educational research sequence work in the OIR? If so, are they paid? If no, explain.

STUDENT INTERVIEW RESPONSE DATA FORM

1. Name______________________ Classification__________

   Address______________________ Department or College __________

   Telephone______________________ Years in this Institution __________

2. Have you had courses in educational research?______________

   Name of the courses______________________________

   Were the courses required?__________ elective?__________

   If required, by whom?______________________________

   What textbook did you study?____________________________

   Did you have a choice of textbooks?______________________

   If so, what alternative texts could you use?______________

3. Are research tools and methods provided for you through reference and library orientation?__________

   If so, what tools?

   How often are you provided with reference and orientation?_______

   What topics did you study?______________________________

   What reference and research books do you particularly study?

4. How is library orientation arranged?

5. How does the reference librarian operate with students?

6. Are bibliographies provided for you?_________ By whom?__________

7. Have you had writing assignments? For what purpose?

8. Have you had oral assignments? To what extent?

9. Name three reference sources which you found to be most helpful in educational research.

   1. ______________________________________

   2. ______________________________________

   3. ______________________________________

10. Have you had a formal computer course?______________

    Who taught the course? Name____________________ Department_______

    Was it a short term course_______ institute_______workshop_______

    semester________ seminar__________ informally taught at computer

    center_________ other___________________.

11. If the above question is yes, did you write your own computer

    programs?________ used "canned computer library programs"?_______

12. Name at least three reference sources which you found to be most helpful in computer application.

    1. ______________________________________

    2. ______________________________________

    3. ______________________________________

13. Were you required to have a prerequisite course(s) in order to take computer application?___________

    If yes, what?

14. Did you have Statistical Analysis as a graduate course?__________

    If yes, was it required?_________________

15. Are there other questions and/or topics on computer application and educational research which you would like to mention?

    1.

    2.

    3.

    4.

    5.

APPENDIX B

RESEARCH CLASSES

Listed below are several dichotomies scaled on the seven-point Likert scale. Rank each dichotomy according to your attitude toward any research classes you have taken at this university. Make a check mark (√) on the scale which best describes your attitude toward the class. For example, if you are given unsuccessful and successful as dichotomies, and if you think the class is more successful, you would make your mark closer to the successful scale. (EXAMPLE)

unsuccessful ____ ____ ____ ____ ____ _√_ ____ successful

RESEARCH CLASS								
GOOD	____	____	____	____	____	____	____	BAD
FAIR	____	____	____	____	____	____	____	UNFAIR
TIMELY	____	____	____	____	____	____	____	UNTIMELY
ACTIVE	____	____	____	____	____	____	____	PASSIVE
VARIED	____	____	____	____	____	____	____	REPETITIVE
FAST	____	____	____	____	____	____	____	SLOW
HARD	____	____	____	____	____	____	____	SOFT
SHARP	____	____	____	____	____	____	____	DULL
STRONG	____	____	____	____	____	____	____	WEAK
IMPORTANT	____	____	____	____	____	____	____	UNIMPORTANT
USUAL	____	____	____	____	____	____	____	UNUSUAL

Enter appropriate class numbers:

I have taken Ed._____; Ed._____; Ed._____

A DESIGN FOR ADVANCED PREPARATION

OF EDUCATIONAL RESEARCHERS

A Dissertation

Presented in Two Volumes to

the Graduate Faculty

East Tennessee State University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

---

by

Fon R. Belcher

June 1972

TABLE OF CONTENTS

Volume II

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

The first assumption of this study was that "Existing university research programs for the preparation of school administrators and educational research personnel reflect the philosophy, beliefs, and work-styles of those who plan and operate them" (Vol. I, p. 8). The major implication, resultant from the data presented in Volume I, was that a custom-made educational research program should be based on a student's work-style. Thus, Volume I runs the gamut from an assumption about producing standard types of researchers to the implication that each student should have his own "custom-made" program. Since the necessities of the latter require the resources of the former, a first requirement of Volume II must be the definition of a system or procedure to arrange the use of standard and/or established and available research resources found in universities, into "custom-made" packages to fit the needs and work-styles of individual advanced graduate students of educational administration.

If the usual university programs to prepare school administrators in research reflect the teacher's philosophy and work-style, a proper program for an advanced student should be developed from that student's philosophy and work-style. Such a task will require both the student researcher and his advisor to engage in making a series of decisions.

The student researcher and his advisor will make decisions concerning the development of the student's custom-made program. The advisor must understand the student's work-style, philosophy, preparation, and ability. Based on this insight, the advisor and student can make better decisions concerning an individualized or custom-made program.

As a result of the problem definition in the student's personalized program, the advisor will be able to suggest alternatives in terms of the possible consequences. The student and advisor must answer the question, "To what extent do we, jointly, possess the competency, personal qualities, and resources necessary to implement this alternative?" This requires an analysis and evaluation of all possible alternatives in the macro systems surrounding the student in the university and in the field in which he may choose to go to work and study.

The choice of alternatives must be made and a committee-approved outline of the program secured. This decision could move the student deeper into his old macro system or also move him out of it into another macro system or field situation with another but similar set of interacting components.

The subsystems which make the components of a custom-made program must be compatible. Activities involved in the implementation of a decision include four basic steps: (1) organizing, (2) planning, (3) programming, (4) operating (OPPO). The first is organizing or establishing the formal structure through which work subdivisions are to be arranged, defined, and coordinated for the specific program.

Second is planning, or working out in broad outline, the things that need to be done and the methods for doing them to accomplish the program developed for the students. Third is programming or outlining the components involved in the custom-made program; and, fourth, operating or carrying out the objectives of the student's custom-made program.

In implementing the program, the total environment must be assessed. The research student and his advisors should discover the available favorable factors to reduce the total environment to a sort of educational macro OPPO system, in which the four subsystems can be developed (Figure 1). These subsystems are the student resource area, the student role/goal subsystem, the university resource subsystem, and institutional support and control subsystem. Within each of these subsystems there are interacting variables. In the student subsystem, the student's work-style, philosophy, preparation, and ability are initially used as inputs. The human and material resources of the university subsystem are composed of program components and advisory personnel. This subsystem composes the second group of inputs. The role subsystem is influenced by the student's future role position and goals. Variables within this subsystem are position, time, place, and design. The institutional support and control subsystem is composed of the acceptance, permission, finance, equipment, and approval variables given to the program. These variables are formal elements and possible inputs from the macro system. The purpose then becomes one of the reduction and use of macro elements in the construction of a micro system for each student. A final result of the interaction of these

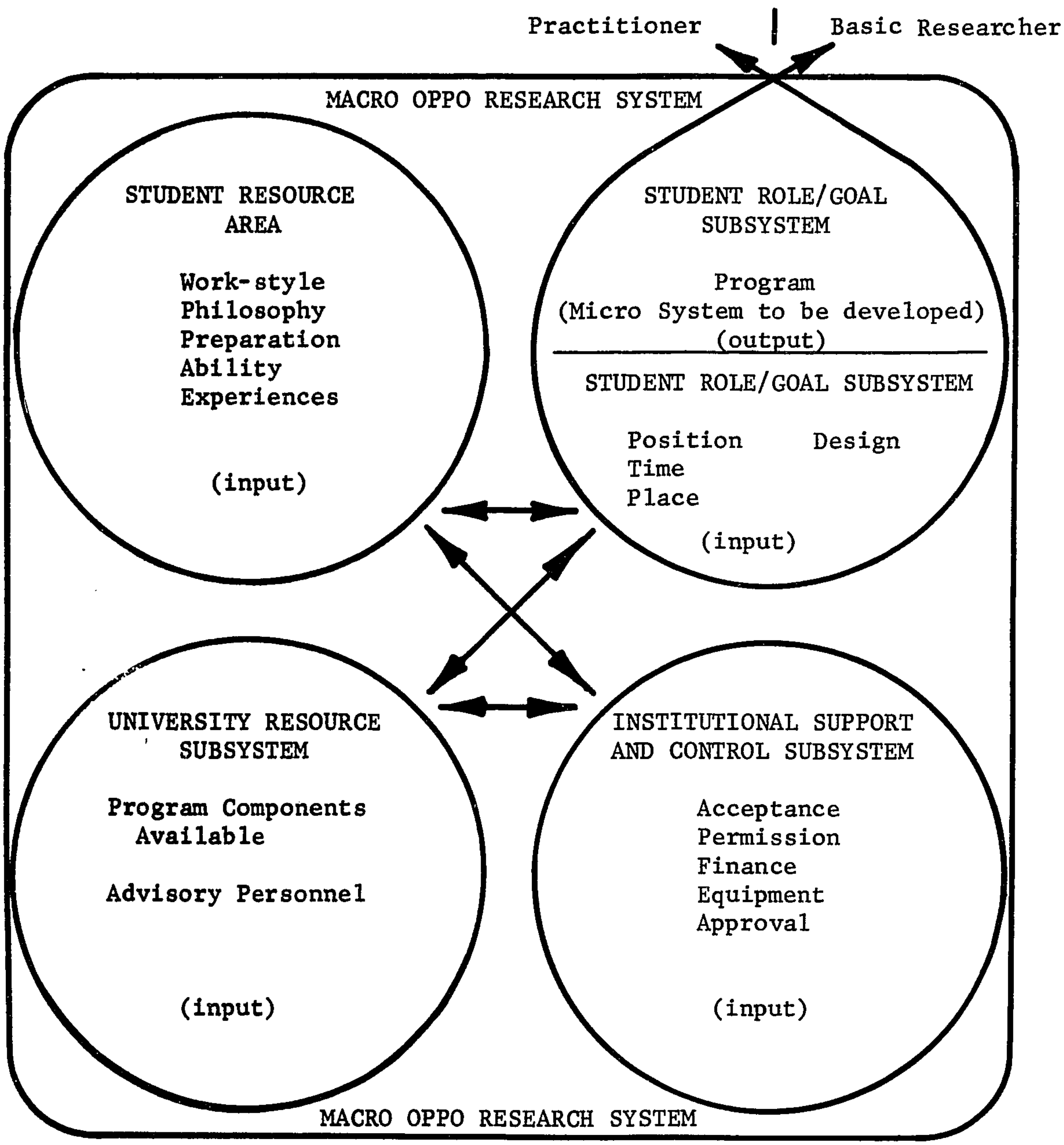

Figure 1

Schema of Interactions in a Macro OPPO Research System
(With Micro System to be Developed)

four subsystems will be reflected in the micro output (either practitioner or basic researcher) of the macro OPPO system.

An examination of the subsystems of Figure 1 shows that administrative decision-making takes into consideration the interaction previously described. This leads into the Micro OPPO Research System: Internship (Figure 2). The Micro OPPO system comprises organizing, planning, programming, and operating in an actual research situation. The nature of the active participants and the available resources in these four interacting subsystems will affect the student and his solution to the problem. For that reason an analysis of these subsystems and a determination of the potential involved is mandatory.

In considering Figures 1 and 2 it is well to keep in mind that Figure 2 represents the development of the upper position of the "Student Role/Goal Subsystem" of Figure 1. The micro OPPO system at the center of Figure 2 could be operational in any macro system, within any type of institution at any time or place. The crux is in the ever-present necessity at all levels to organize, plan, program, and operate. This leads to the consideration of known and unknown variables that surround the student and his advisor as they attempt to reach "custom-made" program decisions.

Acceptance of the concept of a custom-made program for advanced research students makes it impossible for one program to be proposed for every student. Since students have varying work-styles, this will affect the planning of a personal program.

Of the variables listed in Figure 3, some are known and some are unknown. These lists are intended to be suggestive only and may be

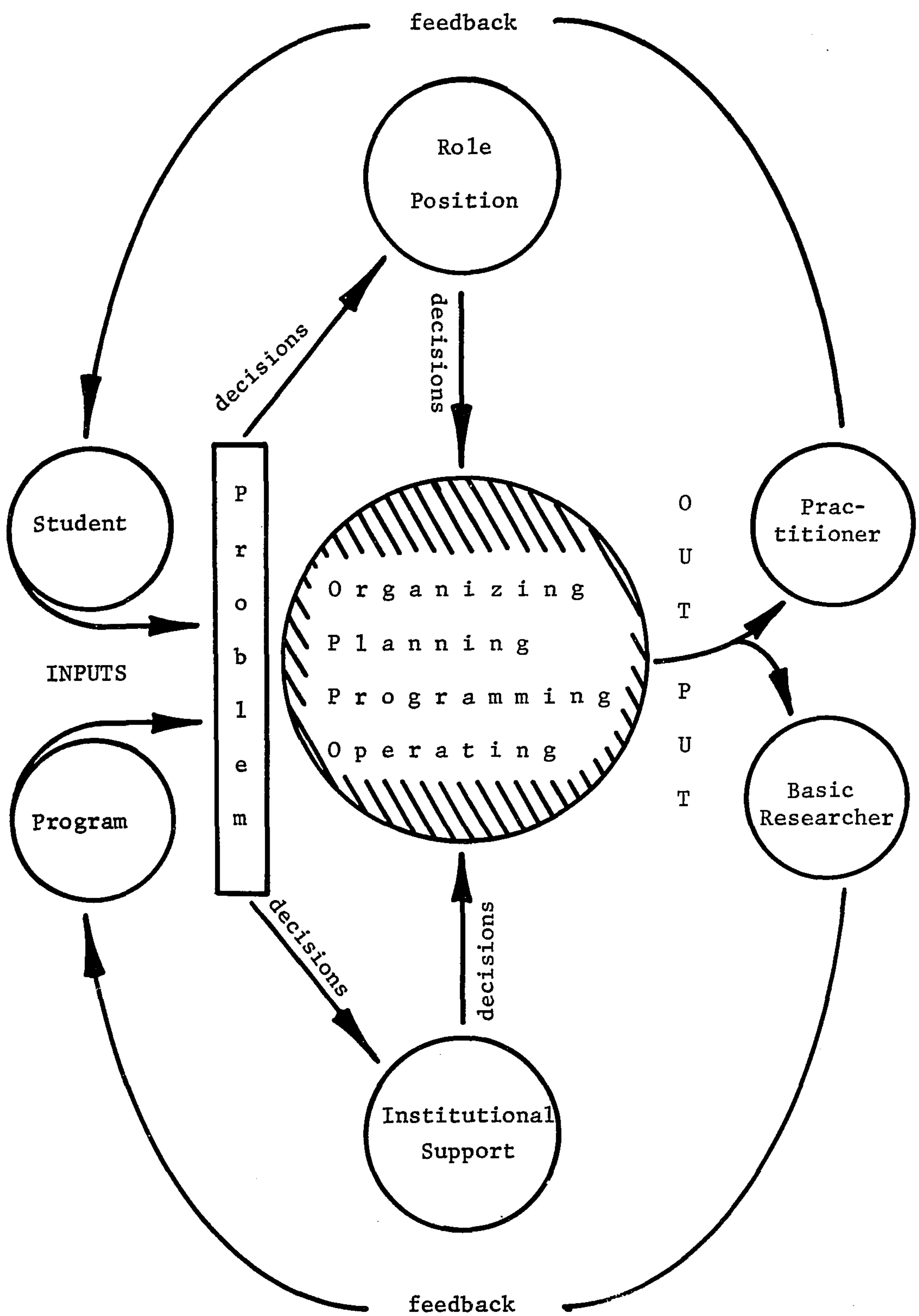

Figure 2

Schema of a Micro OPPO Research System: Internship

Variables (related to)		Known	Unknown
Student	Level of ability		
	Preparation		
	Experience		
	Work-style		
	Role and Goal		
School	Rules and Regulations		
	Offerings		
	Support		
	Attitude		
Resources	Financial		
	Professional		
	Technology		
	Designs and Techniques		
Time and Place	Availability		
	Opportunity		

Figure 3

Schema of Some Known and Unknown Variables
for Consideration in Planning A
Research Preparation Program

extended by students and advisors as specific instances arise that make extension possible or necessary. Figure 3 is intended merely to aid thinking and to help individual research preparation program planners consider the nature and variety of the types of variables involved.

Work-style is philosophically based. As previously stated (p. 1) research programs and research courses dealing with techniques reflect the philosophy and work-styles of specific professors. Matching student work-styles with the work-style resources of professors and technological resources of schools lies in a matrix of interacting variables suggested by Figure 3. It is assumed that the closer the match of student preferences with available professional assistance and institutional resources the greater the prospects for research perfectability at any level and involving any type of work-style or research technique. Diagnosis and honest reflection upon the nature, strength, persistence, recurrence, and resultant evidence of these variables is essential to a good and proper match.

Since this matching involves both the problem and the purpose of the process, it is necessary to consider here a possible methodology for decision-making in terms of a student's program needs. The following schema, (1) Figure 4, describing philosophic viewpoints and characteristics, and (2) Figure 5, suggested as a method for determining work-styles are offered here as guides to promote thinking and the production of reliable devices for measuring and/or estimating the chances of student success in an individually-planned program created for him from the suggested components set forth in this study.

It is presumed that all students need survey-type course work of a basic nature to familiarize them with known and available research

Philosophically-Based Work Styles

<u>The Mathematically Oriented</u> Believes in quantitative approach Work-style would consist of analysis, comparisons, computer Style would be self-directed	<u>An Existentialist</u> Wants to know Why Where to Where from
<u>A Survey Man</u> Would be interested in How many Who When What - Style other-directed	<u>A Pragmatist</u> Interested in experiments Style would be inner-directed
<u>The Planner</u> One who has to be a generalist or an architect Style is nearest to needing or recognizing all styles	<u>A Realist</u> Involves the styles of the Quantitative Approach and the Pragmatist
<u>Those Seeking Status Quo</u> One whose research will try to prove the now	<u>An OPPO ist</u> Is reality oriented Includes all other work-styles

Figure 4

Schema of Work-styles of Eight Candidates
Planning to do Research

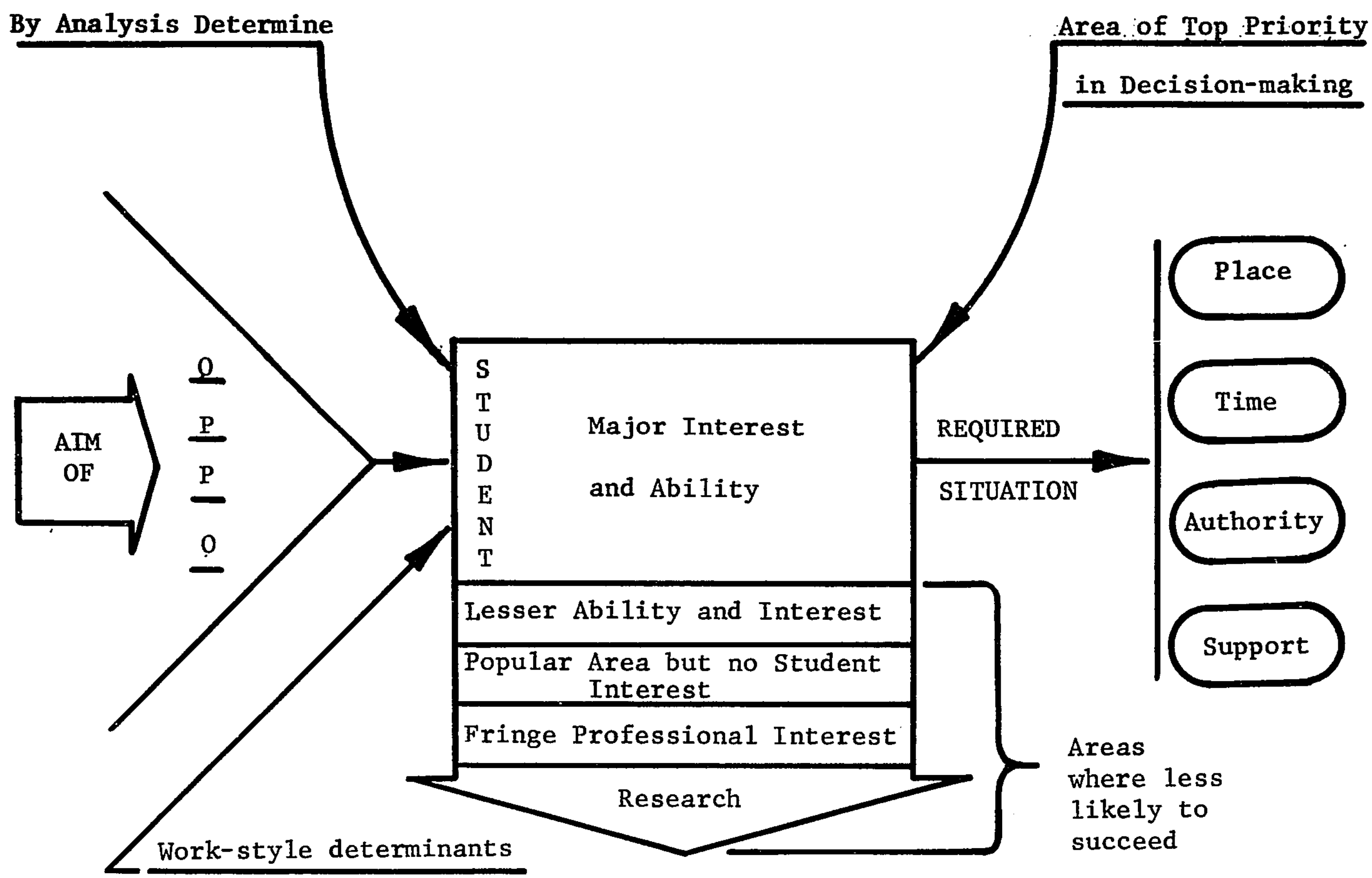

Figure 5

One Schema for Determining Student Work-Style designs, methods, and techniques. It is a major assumption of this study that a specific student has no need for sophisticated techniques until he has first defined a research problem.[1] Hence, a major function of advanced graduate student advisement is to acquaint the student with available program components and to help him see how these research techniques and practices are a necessary and useful means of solving certain types of problems. This process is presented schematically in Figure 5.

The program components which follow are considered representative of certain levels of research competency. Along with familiarity with the accompanying suggested or other current bibliographies, all of the components, or satisfactory evidence of the research competencies they represent (testing to determine), are desirable and should be required of a student electing to prepare himself to be a full-time comprehensive educational researcher. However, the general student of educational administration, destined for a less research-oriented position in terms of his work-style, should be encouraged to choose, along with his advisors, those components necessary to solve his research problem (or problems) after he has become familiar with general research methodology through survey courses.

These components represent the wide scope of preparation for an across-the-board educational researcher, but they need not be taken sequentially. The practitioner should also be allowed to select components appropriate to his needs.

---

[1]Opinion confirmed by Carter V. Good and Ralph Purdy, personal interviews, November 30, 1971, and December 1, 1971.

EVALUATION OF RESEARCH DESIGNS

Table 1 provides a basis on which to evaluate the major advantages and disadvantages of several research designs. A minus sign indicates that the factor represented in the column heading introduces a particular weakness into the design. A plus sign under the column heading indicates that the design controls the weakness. The meaning of a question mark is obvious. Where a blank space occurs, the factor is irrelevant to that design.

An experimental design is said to have high internal validity when control of the factors affecting outcomes is great. When results of the experiment are generalizable to the total universe of the sample, the design has high external validity.

PROGRAM COMPONENTS

In the numbering system which follows, the first digit identifies the research design. Where applicable, the second and third digits identify useful statistical treatments and computer information, in that order. Statistical treatments and computer information suggested here are not intended to be definitive or comprehensive. They are examples of treatments appropriate to the designs they accompany, and, in many cases, to other designs as well. It is anticipated that the student, in consultation with his advisor, will choose appropriate statistical treatments and computer programs. Further information which is helpful in the selection of statistical and computer treatments may be found in the footnotes and bibliographical references accompanying this volume.

Table 1

Major Advantages and Disadvantages of Different Experimental Designs*

	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Comments
Single-Group Designs:								
1. One-shot case study	-	-				-	-	Does not allow for comparison of changes; no premeasures.
2. One-group	-	-	-	-	?	+	+	A type of repeated measurement design but with a single group.
3. Time-series	-	+	+	?	+	+	+	Leads to a complex statistical analysis.
4. Equivalent time samples	+	+	+	+	+	+	+	Generalization is only to other groups which are repeatedly tested.
5. Equivalent materials design	+	+	+	+	+	+	+	Generalization again restricted to groups tested repeatedly.
Separate Control Group Designs:								
6. Static-group	+	?	+	+	+	-	-	Along with designs 1 and 2, this is a pre-experimental design.
7. Pretest-posttest control group design with randomization	+	+	+	+	+	+	+	A true experimental design; generalization restricted to other pretested groups.

*Gilbert Sax, Empirical Foundations of Educational Research (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), p. 381.

Table 1 (continued)

	History	Maturation	Testing	Instrumen-tation	Regression	Selection	Mortality	Comments
8. Solomon four-group design	+	+	+	+	+	+	+	Another true experimental design; requires use of multiple groups.
9. Posttest-only control group design	+	+	+	+	+	+	+	A third true experimental design.
10. Nonequivalent control group design	+	+	+	+	?	+	+	Makes use of intact groups.
11. Separate-sample pretest posttest design	-	-	+	?	+	+	-	Generalization facilitated.

An X represents the exposure of a group to an experimental variable or event, the effects of which are to be measured; 0 refers to some process of observation or measurement.

Pre-experimental Designs

A one-shot case study (1.0). Much research in education today conforms to a design in which a single group is studied only once, subsequent to some agent or treatment presumed to cause change. Such studies might be diagrammed as follows:

X 0

. . . .

In the case studies of Design 1, a carefully studied single instance is implicitly compared with other events casually observed and remembered. The inferences are based upon general expectations of what the data would have been had the X not occurred. Such studies often involve tedious collection of specific detail, careful observation, testing, and the like, and in such instances involve the error of misplaced precision. The study would be much more valuable if the one set of observations were reduced by half and the saved effort directed to the study in equal detail of an appropriate comparison instance. It seems almost unethical at the present time to allow, as theses or dissertations in education, case studies of this nature (i.e., involving a single group observed at one time only). 'Standardized' tests in such case studies provide only very limited help, since the rival sources of difference other than X are so numerous as to render the 'standard' reference group almost useless as a 'control group.' On the same grounds, the many uncontrolled sources of difference between a present case study and potential future ones which might be compared with it are so numerous as to make justification in terms of providing a bench mark for future studies also hopeless. In general, it would be better to apportion the descriptive effort between both sides of an interesting comparison.[2]

Of importance to any experimental design is the blocking and tabling of data.

In most experiments involving the collection and analysis of data, the data are grouped, or blocked, to reduce variability

---

[2]Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally and Company, 1968), pp. 6-7.

and to make the scores more manageable for statistical analysis. The most commonly used methods are to take the mean, median, and mode of several scores. The value obtained is then treated as a single score.

. . . .

If an overall mean were desired, the experimenter would simply add all the correct responses and divide by the number of trials administered.

The median or mode is not frequently used to block raw data. Use of these measures usually occurs only if the data within each set of scores are severely skewed. The median equals the middle score of each block of trials after the scores within each block have been rank ordered. The mode is simply the score that occurs most often.

. . . .

In addition to grouping, it is usually desirable to obtain some idea of the dispersion of the scores. The simplest such measure is the range. This value is computed very simply since it is equal to the difference between the largest and smallest scores in the distribution.

The usefulness of the range is limited. By far the most common and useful measure of dispersion is the standard deviation (s.d.). There are two ways to compute the standard deviation: directly, using difference scores; and indirectly, using the computational formula.[3]

<u>The one-group pretest-posttest design (2.0)</u>. This is used as a 'bad example' to illustrate several of the confounded extraneous variables that can jeopardize internal validity. These variables offer plausible hypotheses explaining an $0_1 - 0_2$ difference, rather than the hypothesis that X caused the difference:

$$0_1 \quad X \quad 0_2$$

The first of these uncontrolled hypotheses is history. Between 0_1 and 0_2 many other change-producing events may have occurred in addition to the experimenter's X. If the pretest (0_1) and the posttest (0_2) are made on different days, then the events in between may have caused the difference.

. . . .

History becomes a more plausible rival explanation of change the longer the $0_1 - 0_2$ time lapse, and might be regarded as a trivial problem in an experiment completed within a one- or two-hour period, although even here, extraneous sources such as laughter, distracting events, et cetera, are to be looked for. Relevant to the variable, history is the feature of

---

[3]James L. Bruning and B. L. Kintz, <u>Computational Handbook of Statistics</u> (Glenview, Illinois: Scott, Foresman and Company, 1968), pp. 2-4.

experimental isolation, which can so nearly be achieved in many physical science laboratories as to render Design 2 acceptable for much of their research.

. . . .

A second rival variable, or class of variables, is designated maturation. This term is used to cover all of those biological or psychological processes which systematically vary with the passage of time, independent of specific external events. Thus between 0_1 and 0_2 the students may have grown older, hungrier, more tired, more bored, et cetera, and the obtained difference may reflect this process rather than X . . . .

A third confounded rival explanation is the effect of testing; the effect of the pretest itself. On achievement and intelligence tests, students taking the test for a second time, or taking an alternate form of the test usually do better than those taking the test for the first time.

. . . .

Instrumentation or 'instrument decay' (Campbell, 1957) is the term used to indicate a fourth uncontrolled rival hypothesis. This term refers to autonomous changes in the measuring instrument which might account for an $0_1 - 0_2$ difference.

A fifth confounded variable in some instances of Design 2 is statistical regression. If, for example, in a remediation experiment, students are picked for a special experimental treatment because they do particularly poorly on an achievement test (which becomes for them the 0_1), then on a subsequent testing using a parallel form or repeating the same test, 0_2 for this group will almost surely average higher than did 0_1. This dependable result is not due to any genuine effect of X, or any test-retest practice effect. It is rather a tautological aspect of the imperfect correlation between 0_1 and 0_2.[4]

<u>The t-test for a difference between a sample mean and the population mean (2.1)</u>. One of the most commonly used tests for significance is the t-test. One use of this test assumes that the mean for some population is known. Knowing this population value, the experimenter can then determine whether the sample he has chosen is significantly different from the population mean.[5]

<u>The t-test for a difference between two independent means (2.2)</u>. Probably the most common use of the t-test is to determine whether the performance difference between two

---

[4]Campbell and Stanley, op. cit., pp. 7-10.

[5]Bruning and Kintz, op. cit., p. 7.

groups of subjects is significant. In most experimental situations, the subjects are randomly assigned to the two groups; one of the groups is manipulated experimentally and the effects of this manipulation are analyzed by comparing the performance of the two groups. There are many instances where the groups are already constituted (e.g., males versus females) and the experimenter wishes to determine whether they differ with respect to some other variable (e.g., height, weight, etc.).[6]

The t-test for related measures (2.3). The t-test is sometimes used to determine the significance of a difference between two correlated means. It is most commonly used in this way when two scores are recorded for the same individuals. The pretest posttest scores would be examples. The second use of this test is in the instance where pairs of subjects in two different groups are "matched" on the basis of some variable to ensure that the pairs of subjects in each group are the same before experimental manipulations are begun. The experimental groups will have the same number of measures since they represent two measures of the same subjects, or matched pairs of subjects. If two measures are taken on the same subjects, the treatments by subject analysis are usually more appropriate.[7]

Completely randomized design: analysis of variance (2.4). This design is basically an extension of the t-test to experiments involving three or more groups. This design would typically be used, for example, if an experimenter were interested in determining reactions to several different drugs. In this case, subjects would be randomly assigned to the several groups and a different drug administered to each group. After the performance of subjects in all groups has been measured, statistical analyses of the data would be undertaken to determine the differential effects of the drugs.[8]

---

[6]Ibid., p. 9.

[7]Ibid., pp. 12-13.

[8]Ibid., p. 17.

The static-group comparison (3.0). The third pre-experimental design needed for development of invalidating factors is the static-group comparison. This is a design in which a group which has experienced X is compared with one which has not, for the purpose of establishing the effect of X.

$$\frac{X - - O_1}{O_2}$$

Instances of this kind of research include, for example, the comparison of school systems which require the bachelor's degree of teachers (the X) with those which do not; the comparison of students in classes given speed-reading training with those not given it; the comparison of those who heard a certain TV program with those who did not, and so on. In marked contrast with the 'true' experiment of Design 6 there are in these Design 3 instances no formal means of certifying that the groups would have been equivalent had it not been for the X. This absence, indicated in the diagram by the dashed lines separating the two groups, provides the next factor needing control, i.e., selection. If O_1 and O_2 differ, this difference could well have come about through the differential recruitment of persons making up the groups; the groups might have differed anyway, without the occurrence of X.

A final confounded variable for the present list can be called experimental mortality, or the production of $O_1 -- O_2$ differences in groups due to the differential drop-out of persons from the groups. Thus, even if in Design 3 the two groups had once been identical, they might differ now not because of any change on the part of individual members, but rather because of the selective dropout of persons from one of the groups.[9]

DATA PREPARATION AND COMPUTER OPERATION

In analysis of data, regardless of research design, the computer, which is appropriate for use with any statistical treatment, is such a vital tool that it is essential that the student know how to prepare his data for electronic processing, and understand the range and scope of computer operation.

---

[9]Campbell and Stanley, op. cit., p. 12.

It should be understood that computer techniques suggested in this study will fit the IBM 1130 computer system, but not necessarily the available program subroutines and the format of input layout of the data. The programmer must take into consideration the program subroutine, its inputs and outputs, relative to the data generated by a specific research study. For that reason, information related to computer programming is included at this point. Wherever these suggested programs are to be used, the researcher and the program analyst should arrive at a compatible data and program organization prior to the beginning of any project involving the IBM 1130 computer.

Preparation of Data

It is assumed that the user is acquainted with the card punch machine (keypunch) and that he can arrange for the posting of his data to cards. This discussion is restricted to consideration of card layouts required and/or recommended.

In general, it is good practice to adopt a standard ten column identification field with column nine reserved for a deck number and column ten reserved for card number. Columns one through eight would usually contain a unique subject identification number and other information upon which sorts of the data files are likely to be based. This information would be repeated on each card of a multiple-card file.

Variable fields should be punched without decimals if a program requiring the 1130 FORMAT Card is to be used. The 1130 FORMAT Card permits insertion of decimals in variables. Some programs offer a selection from several fixed input data layouts. The user may wish to arrange variables or scores in five-column fields with decimal points punched where necessary. It is also possible to arrange input data in uniform fields of one, two, three, or four columns each for these same programs. The main concept to keep in mind when preparing data for analysis by the programs not requiring the 1130 FORMAT Card is uniform field width. A mixture of field widths is not suitable for input to such programs.

The user interested in a more detailed discussion of punch card processing and data card layout should see Chapter 3 in Veldman, D. J., FORTRAN Programming for the Behavioral Sciences.[10]

The Control Cards[11]

There are two kinds of control cards: system control cards and program control cards.

System control cards. //JOB T Card. The card is punched as follows:

Column	Entry
1-2	//
3	Blank
4-6	JOB
7	Blank
8	T
9-10	Blank
11-14	0001
15	Blank
16-19	0002

//XEQ Card is punched as follows:

Column	Entry
1-2	//
3	Blank
4-6	XEQ
7	Blank
8-9	ON

Account Card is punched as follows:

Column	Entry
1	. (period)
2-6	Account Number
7-10	Blank

---

[10] James H. Hogge and Judith H. Picklesimer, "Peabody Statistical Library User's Manual" (Nashville, Tennessee: George Peabody College, 1969), p. 2. (Mimeographed.)

[11] Ibid., pp. 3-4.

Column	Entry
11-25	Name, Problem Title, etc. Information punched in these columns will be printed in the output.
26-30	Blank
31-34	Code Word

//XEQ OFF Card is punched as follows:

Column	Entry
1-2	//
3	Blank
4-6	XEQ
7	Blank
8-10	OFF

Program control cards. Parameter-title card. This card contains information needed for the successful execution of a specific program. Unless otherwise noted, all entries in parameter fields should be right-justified. This means, for example, that if columns five through eight are allocated for a parameter and the actual number to be punched is 37, the number should be punched so that the 7 is in column eight. Leading blanks (in this example, columns five and six) are treated as zeros.

1130 FORMAT card. This card describes the layout (format) of input data for the computer. Columns one through thirteen of this card are always the same:

Column	Entry
1	Blank
2-5	1130
6	Blank
7-12	FORMAT
13	Blank

The remainder of the 1130 FORMAT Card varies according to the characteristics of the program and input data under consideration. Four notations are used to describe data layouts for the programs.

X notation is used to indicate that certain columns of each input data card file are to be skipped. The form of X notation is symbolized as

nX

where n is a number greater than 0 and equal to the number of columns to be skipped. The following examples illustrate X notation:

Specification	Interpretation
6X	skip 6 columns
41X	skip 41 columns
1X	skip 1 column

The use of X notation is appropriate with any program requiring the 1130 FORMAT Card. This flexible notation describes variable (score data) fields. The form of F notation is symbolized as

rFw.d

w = width of the variable field in columns.
d = number of digits after the decimal.
r = number of similar fields in succession.[12]

Consider the following examples:

Specification	Interpretation
6F10.4	6 fields in succession, each 10 columns wide with 4 digits after the decimal.
4F2.0	4 fields in succession, each 2 columns wide with no decimal places.
F5.1	1 field five columns wide with one digit after the decimal.

The decimal point need not actually appear in the variable field. Instead, the 'd' specification may be used to insert a decimal in variable fields where none is punched. The following illustrates how the digits 73145 may be read using F notation:

Specification	Resulting Number
F5.0	73145.0
F5.1	7314.5
F5.2	731.45
F5.3	73.145
F5.4	7.3145
F5.5	.73145

If a decimal point is actually punched in a variable field it will override the F notation field specification.

---

[12]Ibid., pp. 4-5.

For example, if 31.63 were read by F5.0, the number actually read would be 31.63.

I notation is required by some of the programs. It takes the form

rIw

where

w = field width in columns.
r = number of similar fields in succession.

Note that only whole numbers may be read by I notation, since there is no provision for specifying decimal places. The following examples illustrate the use of I notation:

Specification	Interpretation
6I3	6 fields, 3 columns wide.
I7	1 field, 7 columns wide.
30I1	30 fields, 1 column wide.

A notation is required by a few of the programs. It takes the form:

rAw

where

w = field width in columns.
r = number of similar fields in succession.

This notation is primarily used to describe the identification field for each subject (where the program requires input of subject identification).[13]

The following examples suggest how A notation is used:

Specification	Interpretation
A6	1 field, 6 columns wide.
20A1	20 fields, 1 column wide.
A3	1 field, 3 columns wide.

The appropriate field specifications are separated by commas and are enclosed in parentheses, as illustrated in the following example. Consider the following data card layout:

Column	Entry
1-8	Subject identification
9	Deck Code
10	Card Number
11-13	Variable 1
14-16	Variable 2

[13]Ibid., pp. 5-6.

Column	Entry
17-18	Variable 3
19-22	Variable 4

Suppose the program we are using requires us to skip the subject identification. Our 1130 FORMAT Card might appear as follows: 1130 FORMAT (10X, 2F3.0, F2.0, F4.0)

The right parenthesis must appear in or before column 72; however, there is a way to continue the 1130 FORMAT Card beyond one card.

Multiple cards per subject may be handled two ways. If each card in the subject's file contains the same layout, format specifications describing the first card of the set will be automatically repeated by the computer. The other way to handle this situation is through the use of the slash ('/'). When the computer encounters a slash in the 1130 FORMAT Card, it immediately skips to the next card and begins reading. The 1130 FORMAT Card

1130 FORMAT (10X,3F.0/10X,2F2.0)

would be interpreted as follows:
Skip ten columns, read three one-column fields without decimals; skip to the next card, skip ten columns; read two two-column fields without decimals.[14]

CHI-SQUARE ANALYSIS

In considering the Chi-Square analysis using the IBM 1130 computer, the following specific considerations should be observed.

General Description[15]

a. Contingency tables are formed for input data consisting of up to ten categories coded with the integers 0 through 9. Groups of subjects form the rows of each contingency table, and the integer codes form the columns of each table. Combination of categories (columns) may be specified, and hypothesized percentages for a 1 x K table (where K = number of categories) may be tested.
Chi-Square, degrees of freedom, and probability of chance occurrence are computed for each table.

---

[14]Ibid., pp. 6-7.

[15]Ibid., pp. 33-34.

b. The output consists of (for each input variable):
   (1) Observed frequencies
   (2) Expected frequencies
   (3) Observed frequencies minus expected frequencies
   (4) Cell contributions to total chi-square
   (5) Chi-Square
   (6) Degrees of freedom
   (7) Probability of chance occurrence
c. Limitations:
   (1) The number of variables (tables to be formed) may not exceed 20.
   (2) The number of groups (rows of each table) may not exceed 10.
   (3) The number of subjects in each group may not exceed 32,767.
   (4) The number of integer-coded categories of each variable may not exceed 10 (including zero).
   (5) The number of cards per subject may not exceed 5.
d. Estimation of running time:
   Number of seconds = 10 + CV/10 + GV/100 where
   C = number of data cards.
   V = number of variables.
   G = number of groups.

Order of Cards in Job Deck[16]

a. //JOB T Card
b. //XEQ ON Card
c. Account Card
d. //XEQ (name of subroutine)
e. Parameter-Title Card
f. Variable Location Card(s)
g. Missing Data Signal Card (optional)
h. Combination Control Card(s) (optional)
i. Group Control Card
j. Data Cards

   .
   .
   .
   Repeat i and j for each group
   .
   .
   .

k. Blank Card (stops routine)
l. //XEQ OFF
m. Blank Card

   Note: Items e through j may be repeated for additional data sets as desired.

---

[16]Ibid., p. 34.

Card Preparation[17]

d. //XEQ
This card is punched as follows:

Column	Entry
1-2	//
3	Blank
4-5	XEQ
7	Blank
8-10	(name of subroutine)

e. Parameter-Title Card.

Col. 1-2	Number of variables (tables to be formed).
Col. 3-4	Number of groups (rows of tables).
Col. 5-6	Number of Combination Control Cards to be read.
Col. 7	Number of cards per subject. If a one (1) is punched in this column, a Missing Data Signal Card follows the Variable Location Card(s). If this column is blank or contains a zero, zero codes will be treated as missing data (i.e., excluded from all analyses).
Col. 9-80	Alphanumeric Title. These columns must not be left completely blank or premature termination will occur.

f. Variable Location Card(s).
Each of these cards corresponds to one of the data cards in a subject's card set. In other words, if the data cards include three cards per subject, three Variable Location Cards are required.
There is a one-for-one correspondence between Variable Location Card columns and data card columns. A 1 (one) is punched in each Variable Location Card column which corresponds to a data card column in which a variable's codes are punched. Other Variable Location Card columns are left blank. There must, therefore, be V ones punched in the Variable Location Card columns (where V = number of variables).

---

[17]Ibid., pp. 34-35.

If each subject's card set includes one or more cards from which no codes are to be read for a given computer run, there must nevertheless exist a Variable Location Card for each card of the subject card set. In this case, one or more Variable Location Cards would be completely blank.[18]

g. Missing Data Signal Card (optional).
Each column of this card corresponds to an input variable (in serial order of input). If a one (1) is punched in a column of this card, the table formed for the corresponding variable will include zero codes. If a column of this card is left blank or contains a zero, the table formed for the corresponding variable will exclude zero codes. The Missing Data Signal Card is omitted when column 8 of the Parameter-Title Card is blank or contains a zero.[19]

h. Combination Control Card(s) (optional).
The Combination Control Card (one for each variable for which categories are to be combined and/or hypothesized percentages specified) permits lumping of category codes and/or specification of hypothesized percentages for a one row (group) by k categories table.[20]

ADDITIONAL PROGRAM COMPONENTS

True Experimental Designs

The pretest-posttest control group design (4.0). Design 4 uses equivalent groups achieved by randomization. Design 4 takes this form:

$$\begin{array}{llll} R & O_1 & X & O_2 \\ R & O_3 & & O_4 \end{array}$$

History is controlled insofar as general historical events that might have produced an O_1--O_2 difference would also produce an O_3--O_4 difference. Note, however, that many supposed utilizations of Design 4 (or 5 or 6) do not control for unique intrasession history. Furthermore, the typical experiment in the Journal of Experimental Psychology does achieve control of intrasession history through testing students and animals individually and through assigning the students and experimental periods at random to experimental or control conditions . . . .

---

[18]Ibid., p. 35.

[19]Ibid.

[20]Ibid.

The optimal solution is a randomization of experimental occasions, with such restrictions as are required to achieve balanced representation of such highly likely sources of bias as experimenters, time of day, day of week, portion of semester, nearness to examinations, et cetera. The common expedient of running experimental subjects in small groups rather than individually is inadmissible if this grouping is disregarded in the statistical analysis. . . .

Maturation and testing are controlled in that they should be manifested equally in experimental and control groups. Instrumentation is easily controlled where the conditions for the control of intrasession history are met, particularly where the 0 is achieved by student responses to a fixed instrument such as a printed test. Where observers or interviewers are used, however, the problem becomes more serious. If observers are few enough not to be randomly assignable to the observation of single sessions, then not only should each observer be used for both experimental and control sessions, but in addition, the observers should be kept ignorant of which students are receiving which treatments, so that the knowledge will not bias their ratings or records . . . .

Regression is controlled as far as mean differences are concerned, no matter how extreme the group is on pretest scores, if both experimental and control groups are randomly assigned from the pool of the same extremes. . . .

Selection is ruled out as an explanation of the difference to the extent that randomization has assured group equality at time R. . . .

The data made available by Design 4 make it possible to tell whether mortality offers a plausible explanation of the 0_1--0_2 gain. . . .

. . . .

The factors of internal invalidity which have been described so far have been factors which directly affect 0 scores. . . . The threats to external validity, on the other hand, can be called interaction effects, involving X and some other variable. They thus represent a potential specificity of the effects of X to some undesirably limited set of conditions. . . .

. . . .

Interaction of testing and X. In discussion of experimental design per se, the threat of the pretest to external validity was first presented by Solomon (1949). . . .

. . . .

The effect of the pretest upon X as it restricts external validity is of course a function of the extent to which such repeated measurements are characteristic of the universe to which one wants to generalize. . . .

Interaction of selection and X. While Design 4 controls for the effects of selection at the level of explaining away experimental and control group differences, there remains the possibility that the effects validly demonstrated hold only for that unique population from which the experimental and control groups were jointly selected . . . .

. . . .

Other interactions with X. In parallel fashion, the interaction of X with the other factors can be examined as threats to external validity. Differential mortality would be a product of X rather than interactive with it. . . .[21]

Educational researchers are often confronted with samples involving one or more predictor scores. Using these predictors one should be able to account for much of the individual difference variance (subjects within groups variance). A procedure for making use of predictor scores to reduce the estimate of random error is provided by using the analysis of covariance.

Suppose K1 and K2 are two treatment groups (Figure 6).[22] The scatterplot of their dependent variable scores (Y) and their predictor scores (X) shows that the correlation between the two is high.

If the predictor measure is ignored the variability within the K1 and K2 distributions is seen to be quite large. If, however, instead of basing the variance estimator for random error on the deviations of the Y scores from the Y mean, it is based on the deviations from the Y-on-X regression line, then the estimate of error variance is much smaller. This latter estimate is called the standard error of estimate. The new null hypothesis is: the regression lines are not significantly far apart when compared with the deviations of scores about the regression lines. Since the regression line is mathematically determined to pass through the mean of both the criterion and the predictor variables, this null is equivalent to the null for the difference between means.

Following are two groups with five pairs of scores each.

K1

X	Y
2	1
3	3
6	5
5	7
8	9

K2

X	Y
2	5
3	7
6	6
5	8
8	10

[21]Campbell and Stanley, op. cit., pp. 13-20.

[22]Laird Heal, "Covariance Notes" (Nashville, Tennessee: George Peabody College, 1969), p. 1. (Mimeographed.)

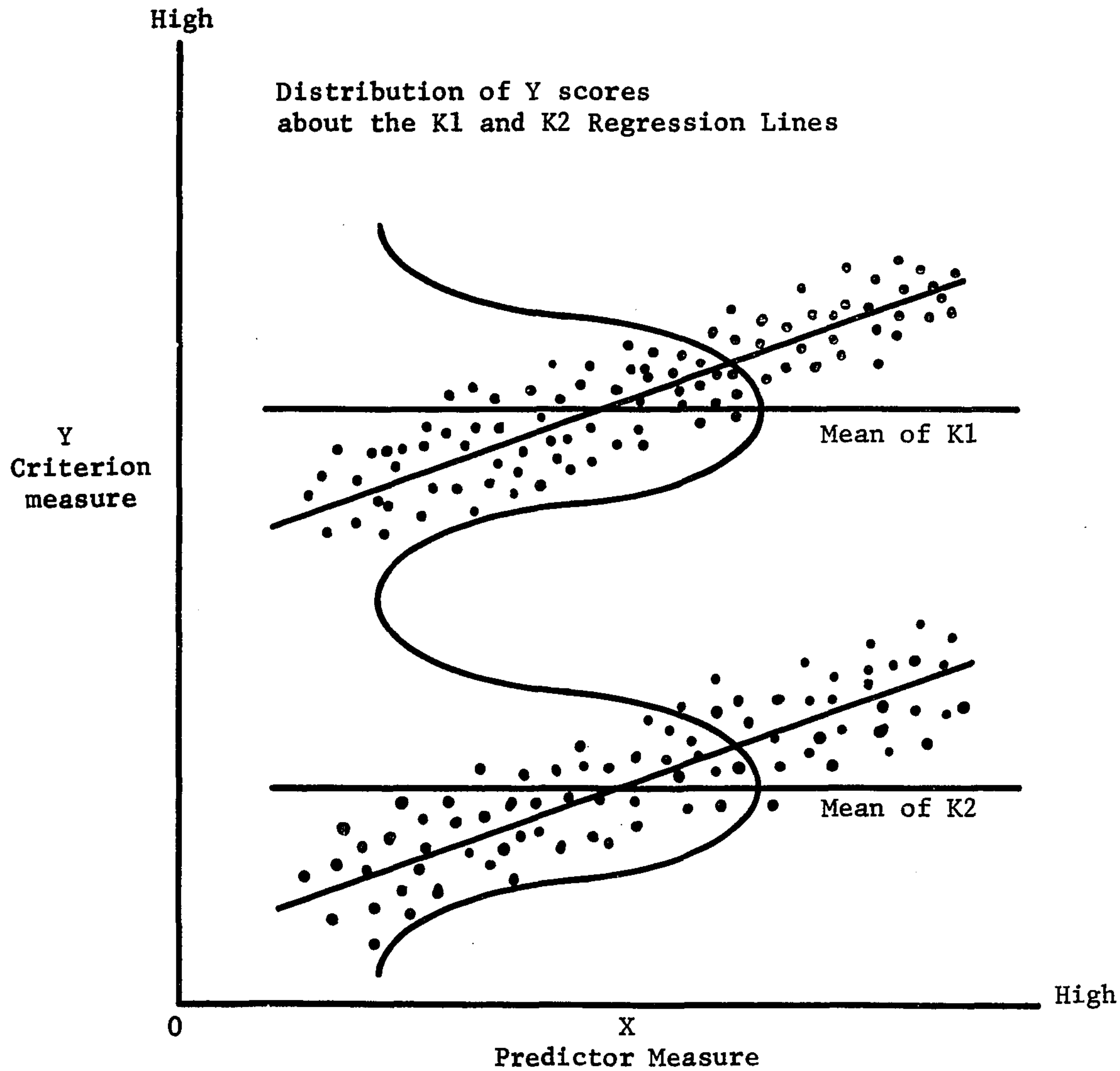

Figure 6

A Scatterplot of Dependent Variables and Predictor Scores
of Two Treatment Groups

The general strategy is to:
(1) calculate the sums of the squared criterion and predictor scores and also the sum of the cross products for each group and for all groups combined;
(2) calculate the predictor and criterion sums for squares and deviation cross products for the total set of scores and for the scores within each group;
(3) calculate the pooled within-groups regression coefficient;
(4) calculate the adjusted total sum of squares, the adjusted between groups sum of squares;
(5) calculate the adjusted means.[23]

1. ΣX, ΣY, ΣX^2, ΣY^2, ΣXY for each cell and for all cells

	ΣX	ΣY	ΣX^2	ΣY^2	ΣXY
K1	24	25	138	165	148
K2	24	36	138	274	187
Total	48	61	276	439	335

2a. ΣX^2, ΣY^2, ΣXY (Total)

$$\Sigma X^2_{Tot} = \Sigma (X - \bar{\bar{X}})^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{nk} \qquad k = \text{no. of groups}$$

$$\Sigma Y^2_{Tot} = \Sigma (Y - \bar{\bar{Y}})^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{nk}$$

$$\Sigma XY_{Tot} = \Sigma (XY - \bar{\bar{XY}}) = \Sigma XY - \frac{(\Sigma X)\ (\Sigma Y)}{nk}$$

$$\Sigma X^2_{Tot} = 276 - \frac{(48)^2}{(5)(2)} = 45.6$$

$$\Sigma Y^2_{Tot} = 439 - \frac{(61)^2}{(5)(2)} = 66.9$$

$$\Sigma XY_{Tot} = 335 - \frac{(48)(61)}{(5)(2)} = 42.2$$

---

[23]Ibid., pp. 1-2.

2b. ΣX^2, ΣY^2, ΣXY (within)

$$\Sigma X^2_w = \Sigma X^2 - \left(\frac{(\Sigma X)^2_{k1} + (\Sigma X)^2_{k2} + \ldots}{n \text{ per cell}}\right)$$

$$\Sigma Y^2_w = \Sigma Y^2 - \left(\frac{(\Sigma Y)^2_{k1} + (\Sigma Y)^2_{k2}}{n \text{ per cell}}\right)$$

$$\Sigma XY_w = \Sigma XY - \left(\frac{(\Sigma X)\ (\Sigma Y)_{k1} + (\Sigma X)\ (\Sigma Y)_{k2}}{n \text{ per cell}}\right)$$

$$\Sigma X^2_w = 276 - \frac{(24)^2 + (24)^2}{5} = 45.6$$

$$\Sigma Y^2_w = 439 - \frac{(25)^2 + (36)^2}{5} = 54.8$$

$$\Sigma XY_w = 335 - \frac{(24)(25) + (24)(36)}{5} = 42.2$$

3a. $bYX_{Tot} = \frac{\Sigma XY_T}{\Sigma X^2_T} = \frac{42.2}{45.6} = .925$

3b. $bYX_w = \frac{\Sigma XY_w}{\Sigma X^2_w} = \frac{42.2}{45.6} = .925$

4a. $SS_{Tot} - adj = \Sigma Y^2_T - (b_T)\ (\Sigma XY_T)$

$SS_{Tot} - adj = 66.9 - (.925)(42.2) = 27.86$

4b. $SS_{w-adj} = \Sigma Y^2_w - (b_w)\ (\Sigma XY_w)$

$= 54.8 - (.925)\ (42.2)$

$= 15.76$

4c. $SS_B - adj = SS_{Tot} - adj - SS_w - adj$

$= 27.86 - 15.76$

$= 12.10$

The analysis of covariance summary is as follows:[24]

Source	df	SS
Groups	1	12.10
Sbs/Gps(adj)	7	15.76
Total	8	27.86
Pooled within groups regression	(1)	(39.04)

<u>The Solomon Four-Group Design (5.0)</u>. While Design 4 is more used, Design 5, the Solomon (1949) Four-Group Design, deservedly has higher prestige and represents the first explicit consideration of external validity factors. The design is as follows:

$$R \quad O_1 \quad X \quad O_2$$
$$R \quad O_3 \quad \quad O_4$$
$$R \quad \quad X \quad O_5$$
$$R \quad \quad \quad O_6$$

By paralleling the Design 4 elements (O_1 through O_4) with experimental and control groups lacking the pretest, both the main effects of testing and the interaction of testing and X are determinable. In this way, not only is generalizability increased, but in addition, the effect of X is replicated in four different fashions: $O_2 \quad O_1$, $O_2 \quad O_4$, $O_5 \quad O_6$, and $O_5 \quad O_3$. The actual instabilities of experimentation are such that if these comparisons are in agreement, the strength of the inference is greatly increased. Another indirect contribution to the generalizability of experimental findings is also made, in that through experience with Design 5 in any given research area one learns the general likelihood of testing-by-X interactions, and thus is better able to interpret past and future Design 4s. In a similar way, one can note (by comparison of O_6 with O_1 and O_3) a combined effect of maturation and history.[25]

[24]Ibid., pp. 3-5.

[25]Campbell and Stanley, op. cit., pp. 24-25.

<u>Analysis of variance (5.1)</u>. The analysis of variance refers to a means of testing whether or not two sample variabilities differ. Variance is an index of variability. It is the square of the standard deviation of a population of scores. (MS) Mean square is a sample variance. It is an unbiased estimator of the population from which the sample was drawn. Computationally, it is the sum of the squared deviations of scores about their mean divided by one less than the number of these scores:

i.e.

$$MS = \frac{\Sigma(X - \overline{X})^2}{N - 1} = \frac{SS}{df} = \frac{\text{Sum of Squares}}{\text{degrees of freedom}}$$

F-ratio is a statistic. It is the ratio of two mean squares of samples drawn randomly and independently from the same normally distributed population.

It is distributed with a mean of approximately 1.0 and is skewed to the right. There are really a family of these distributions. The F distribution that fits a particular pair of mean squares depend upon the degrees of freedom of both the numerator MS and the denominator MS.

If the F-ratio is unusually large, having less than five percent chance of occurring under the assumptions implied above, then the decision is usually made to reject the null; i.e. The decision is made that the mean squares came from different populations.

Application to test differences in sample variances: One way to apply the F-test is to compare the variance estimators from two samples of scores to test the null hypothesis that these two mean squares came from the same population:

e.g.

Sample 1	Sample 2
1	1
2	5
3	9
6	15

Sample 1:

SS = Sum of squared deviations of scores about the mean

$$= \Sigma(X - \overline{X})^2 = \Sigma X^2 - \frac{(\Sigma X)^2}{n}$$

$$SS_1 = (1)^2 + (2)^2 + (3)^2 - \frac{(6)^2}{3} = 14 - 12 = 2$$

$$SS_2 = (1)^2 + (5)^2 + (9)^2 - \frac{(15)^2}{3} = 107 - 75 = 32$$

df = degrees of freedom = number of scores minus 1

$df_1 = 3 - 1 = 2$

$df_2 = 3 - 1 = 2$

$$MS_1 = \frac{SS_1}{df_1} = \frac{2}{2} = 1$$

$$MS_2 = \frac{SS_2}{df_2} = \frac{32}{2} = 16$$

It is conventional to put the numerically large MS in the numerator of the F-ratio.

$$F = \frac{MS_2}{MS_1} = \frac{16}{1} = 16$$

Referring to the F table, $F(2,2)_{.05} = 19.0$ Numerator of d.f. indexed at the top of the table. Denomonator d.f. indexed at the side of the table.

These M.S.'s do not differ significantly. If the F had exceeded 19, the null would have been rejected and the decision made to accept the alternative decision that MS_1 and MS_2 were from different populations. Applications to test differences among means: While the preceding example shows one use of the F statistic, it does not illustrate its most common use, which is to make a ratio of two variance estimators that are taken from independent parts of the same set of scores. The ordinary procedure is to separate the variability that can be attributed to interventions or treatments from variability that can be attributed to random fluctuations within treatments: Thus, an F-ratio is constructed:

$$F = \frac{\text{MS Between Treatments}}{\text{MS Within Treatments}}$$

For example, if the prior example is considered as an experiment with two groups, the variability within groups would be pooled from the two groups' M S's as follows:

$$\text{MS pooled within treatments} = \frac{SS_1 + SS_2}{df_1 + df_2}$$

$$= \frac{2 + 32}{2 + 2} = \frac{34}{4}$$

$$= 8.5$$

The SS Between is equal to the sum of squared deviations of the group mean about the grand mean, adjusted for the stability of sampling, or

$$\text{SS Between} = (\text{n per group}) \ \Sigma(\bar{X} - \bar{\bar{X}})^2$$

$$= \frac{\Sigma X^2}{\text{n per gp}} - \frac{(\Sigma X)^2}{N_{Total}} = \frac{(6)^2}{3} + \frac{(15)^2}{3} - \frac{(21)^2}{6}$$

$$= 13.5$$

$$\text{df Between} = \#\ \text{gps} - 1 = 2 - 1 = 1$$

$$\text{MS Between} = \frac{SS_{Between}}{df_{Between}} = \frac{13.5}{1} = 13.5$$

When an F-ratio is needed to test a Between versus Within null hypothesis, it is conventional to have the MS_B in the numerator. The denominator in this context is often called the error term or error variance.

$$F = \frac{MS_B}{MS_W} \quad \frac{13.5}{8.5} = 1.588$$

$$F\ (1,\ 4)_{.05} = 7.71$$

The two groups do not differ significantly. The 2-group F-test (1df in the numerator) is equivalent to a t-test for independent groups. ($F = t^2$)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\dfrac{\Sigma X_1^2 - \dfrac{(\Sigma X_1)^2}{n_1} + \Sigma X_2^2 - \dfrac{(\Sigma X_2)^2}{n_2}}{(n_1 + n_2) - 2}\right]\left[\dfrac{1}{n_1} + \dfrac{1}{n_2}\right]}}$$

$$= \frac{2 - 5}{\sqrt{\left[\dfrac{14 - 12 + 107 - 75}{(3 + 3) - 2}\right] \left[\dfrac{1}{3} + \dfrac{1}{3}\right]}}$$

$$= \frac{-3}{\sqrt{\left(\frac{34}{4}\right)\left(\frac{2}{3}\right)}} =$$

$$F = t^2 = \left(\frac{-3}{\sqrt{\left(\frac{34}{4}\right) \cdot \left(\frac{2}{3}\right)}}\right)^2 = \frac{9}{\frac{34}{6}} = \frac{(6)(9)}{34} = 1.588$$

Total Mean Square: One way to look at the MS_B and MS_W is to consider them as independent parts of the total variance estimator. The MS (total) is the sum of the squared deviations of all scores about the mean.

$$MS_T = \frac{\Sigma\ (X - \bar{\bar{X}})^2}{df_{Total}}$$

$$= (1)^2 + (2)^2 + (3)^2 + (1)^2 + (5)^2 + (9)^2 - \frac{(21)^2}{6}$$

$$= 47.5$$

$$df_T = 6 - 1 = 5$$

In order for MS_B and MS_W to be independent the following must hold:

$$SS_T = SS_B + SS_W$$

$$df_T = df_B = df_W$$

It is often convenient to calculate the SS_T first and then get the SS_W by subtracting the SS_B from it. Similarly

$$df_W = df_T - df_B$$

Thus:

$$MS_W = \frac{SS_T - SS_B}{df_T - df_B}$$

Note that independent mean squares are pooled by adding numerators and denominators separately.

For communication, it is often useful to present a summary of analysis of variance calculations:[26]

Source	df	SS	MS	F
Between	1	13.5	13.5	1.588
Within	4	34.0	8.5	
Total	5	47.5		

Single classification analysis of variance with multiple groups and/or trials (5.1.1).

General Description[27]

a. This program performs analysis of variance with multiple groups and/or trials. The number of subjects need not be the same for all cells, and, if desired, zeros and blanks may be treated as missing data.
b. The output consists of:
   (1) Cell means.
   (2) Analysis of variance table including source, sum of squares, degrees of freedom, mean square, F-ratio, and probability of chance occurrence.
   (3) Number of subjects with complete data (if zeros and blanks are treated as missing data).
c. Limitations:
   (1) The number of dependent variables may not exceed 10.
   (2) The number of groups may not exceed 10.
   (3) The number of trials may not exceed 10.
   (4) The number of subjects in any group may not exceed 32,767.
d. Estimation of running time:
   Number of seconds = 45 + C/2 + 5V where
   C = total number of cards for all subjects.
   V = number of dependent variables.

---

[26] Laird W. Heal, "ANOVA Lecture Notes" (Nashville, Tennessee: George Peabody College, 1969), pp. 1-8. (Mimeographed.)

[27] Hogge and Picklesimer, op. cit., p. 26.

Order of Cards in Job Deck[28]

a. //JOB Card

b. //XEQ ON Card

c. Account Card

d. Program deck including 1130 FORMAT Card

e. Parameter-Title Card

f. Missing Data Signal Card (Optional)

g. Group Control Card

h. Data Cards

(Repeat g. and h. for each group of subjects)

i. //XEQ OFF Card

j. Blank Card

Card Preparation[29]

d. Program Deck including 1130 FORMAT Card.

The program deck should include required system cards except the //JOB Card, //XEQ ON Card, Account Card, and //XEQ OFF Card. The 1130 FORMAT Card is inserted in the program deck after the card which contains the comment: INSERT 1130 FORMAT CARD NEXT.

e. Parameter-Title Card.

Col. 1-2	Number of dependent variables.
Col. 3-4	Number of groups (set equal to 1 if repeated measures analysis of variance on a single group is desired).
Col. 5-6	Number of trials (set equal to 1 if single classification analysis of variance without repeated measures is desired).

[28]Ibid., pp. 26-27.

[29]Ibid., pp. 27-28.

Col. 7	Missing data card signal: 1 if missing data signal card follows Parameter-Title Card, or 0 if inspection for missing data is not desired.
Col. 8-80	Alphanumeric Title.

Note: All entries in parameter fields must be right-justified.

f. Missing Data Signal Card.

Beginning with column 1, each dependent variable is assigned a single-column signal (where the column number corresponds to the dependent variable number):

0 if zero scores and blank fields are to be considered valid scores, or

1 if zero scores and blank fields are to signify missing data.

When multiple trials are part of a design and the missing data signal for a dependent variable = 1, a subject must have valid scores for all trials to be included in the analysis for that dependent variable.

g. Group Control Card.

Col. 1-5	Number of subjects in the group (right-justified).
Col. 6-80	Alphanumeric Group Title.

h. Data Cards.

The sample size may be different from group to group, but the number of dependent variables must be the same for all groups.

References

The present program is based on Veldman's Program ANOVAR in

Veldman, D. J. Fortran Programming for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1967. Pp. 247-257.

The statistical procedure is discussed in

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1962. Pp. 46-138.

Regression analysis for raw data (5.1.2).

General Description

a. This program computes means, sigmas, Pearson product-moment correlations, regression equations, and F tests for comparison of R-squares of selected regression equations.

b. The output consists of:

(1) Means (optional)

(2) Sigmas (optional)

(3) Correlation matrix (optional)

For each regression equation:

(4) Beta weights
(5) Raw score (b) weights
(6) Regression constant
(7) Multiple R
(8) Multiple R-squared
(9) Iteration sequence (optional)

For each F test:

(10) Degrees of Freedom
(11) F-ratio
(12) R-squares of two equations
(13) Probability of chance occurrence of F ratio

c. Limitations:

(1) The number of variables (all criteria and predictors) may not exceed 40.

(2) The number of subjects may not exceed 9,999.

(3) The number of equations may not exceed 35.

d. Estimation of running time:

Number of seconds = $45 + CV/10 + V^2/20 + E + T$ where

C = number of data cards
V = number of variables
E = number of equations
T = number of F tests

Order of Cards in Job Deck

a. //JOB T Card
b. //XEQ ON Card
c. Account Card
d. Program deck including 1130 FORMAT Card
e. Parameter-Title Card
f. Data Cards
g. Model Card(s)
h. F Test Card(s)
i. //XEQ OFF Card
j. Blank Card

Card Preparation

d. Program deck including 1130 FORMAT Card.

The program deck should include all required system cards except the //JOB T Card, //XEQ ON Card, Account Card, and //XEQ OFF Card.

The 1130 FORMAT Card must specify only variable fields in F notation. Use X notation to skip each subject's identification field. The 1130 FORMAT Card is inserted in the program deck after the card which contains the comment: INSERT 1130 FORMAT CARD NEXT.

e. Parameter-Title Card.

Col. 1-2	Number of variables (all criteria and predictors)
Col. 3-6	Number of subjects
Col. 7-8	Number of equations (models)
Col. 9-10	Number of F tests
Col. 11	If a one (1) is punched in this column, means, sigmas, and the correlation matrix will be printed.
Col. 12-80	Alphanumeric Title.

f. Data Cards.

Data are symbolized by X_{ij}, where i refers to subject, and j refers to variable.

The number of variables must be the same for all subjects.

g. Model Card(s).

This card (one for each regression equation) describes a regression equation or 'regression model.'

Col. 1-2	Model number. This entry is the number of the model in its order of appearance.
Col. 3-4	Number of criterion variable (in serial order of input).
Col. 5	If a one (1) is punched in this column, the iteration sequence will be printed.
Col. 6-8	Blank.
Col. 9-10	Number of predictor-group specifications to follow (maximum = 14).
Col. 11	Blank.
Col. 12-15, 17-20, 22-25, etc.	The first two digits designate the first variable of a group and the last two digits designate the last variable of the same group. Note that a blank column appears between predictor-group specifications.

Examples (b = blank, and each example portrays the portion of the model card beginning in column 12):

0203b0507b0909 specifies the following variables as predictors: 2, 3, 5, 6, 7, and 9. In this case, columns 9-10 = 03.

0121 specifies variables 1 through 21 as predictors. In this case, columns 9-10 = 01.

h. F test Card(s).

This card (one for each F test) specifies two regression equations whose R-squares are to be compared in an F test.

Col. 1-5	Model number (i.e., position in order of appearance) of larger equation or 'full model.'
Col. 6-10	Model number of smaller equation or 'restricted model.'
Col. 11-15	Numerator degrees of freedom.
Col. 16-20	Denominator degrees of freedom.
Col. 21-80	Alphanumeric Title.

To test a model against zero (i.e., to test the significance of a single regression equation), leave columns 6-10 blank.

Degrees of Freedom. If

N = number of subjects,
i = number of predictors in the full model or larger equation, and
j = number of predictors in the restricted model or smaller equation,

then, for most problems,

numerator degrees of freedom = i - j, and denominator degrees of freedom = N - i - 1.

When some predictors can be expressed as linear combinations of others, the computation of the degrees of freedom is more complicated. In such cases, see p. 106 of Bottenbert and Ward (below).[30]

References

This program is an adaptation of Program REGRAN in

Veldman, D. J. Fortran Programming for the Behavioral Sciences. New York: Holt, Rinehart, and Winston, 1967. Pp. 295-307.

The statistical technique is discussed in

Bottenbert, R. A., and Ward, J. H., Jr. Applied Multiple Linear Regression. U. S. Department of Commerce Office of Technical Services, AD413128, 1963.

[30] Ibid., pp. 42-45.

The Posttest-only Control Group Design (6.0). While the pretest is a concept deeply embedded in the thinking of research workers in education and psychology, it is not actually essential to true experimental designs. . . . Within the limits of confidence stated by the tests of significance, randomization can suffice without the pretest . . . . Its form is as follows:

$$\begin{matrix} R & X & O_1 \\ R & & O_2 \end{matrix}$$

The design can be considered as the two last groups of the Solomon Four-Group Design, and it can be seen that it controls for testing as main effect and interaction, but unlike Design 5 it does not measure them. However, such measurement is tangential to the central question of whether or not X did have an effect . . . .

In the repeated-testing setting of much educational research, if appropriate antecedent variates are available, they should certainly be used for blocking or leveling, or as covariates. . . .

Many problems exist for which pretests are unavailable, inconvenient, or likely to be reactive, and for such purposes the legitimacy of Design 6 still needs emphasis in many quarters. In addition to studies of the mode of teaching novel subject materials, a large class of instances remains in which (1) the X and posttest O can be delivered to students or groups as a single natural package, and (2) a pretest would be awkward. Such settings frequently occur in research on testing procedures themselves, as in studies of different instructions, different answer-sheet formats, et cetera. Studies of persuasive appeals for volunteering, et cetera are similar. Where student anonymity must be kept, Design 6 is usually the most convenient. In such cases, randomization is handled in the mixed ordering of materials for distribution.[31]

2x2x2 Factorial Analysis of Variance (6.1). An experiment has been done in which 16 males and 16 females are either given a driving course (T) or not (NT), there being 8 in each subgroup. Half of the subjects in each of these subgroups is in the top quartile of a motor coordination test and the other half in the bottom quartile (C vs. UC). The dependent variable was the students' performance on a driver's test that was given at the end of the course (or at the same time, in the case of the nontrained (NT) groups).

[31]Campbell and Stanley, op. cit., pp. 25-26.

	Male		Female	
	C	UC	C	UC
	4	3	7	3
T	4	2	9	2
	1	6	11	5
	7	2	10	1
NT	7	2	3	3
	1	5	4	1
	8	2	4	4
	7	1	2	0

Do an analysis of variance of these data partitioning the between sum of squares (and df) into the traditional 2x2x2 partition (one effect for each of 7 df).

Hint: Find the coefficients for main effects; cross-multiply these sets to find the coefficients for the three double-order interactions and then multiply all three sets together to find the coefficients.

Evaluate the status of each of the 7 hypotheses implied by the design. The analysis of this design proceeds in the same way as that for a simple randomized groups factorial design. There are eight groups, making several df for treatments. Because each treatment factor was two levels, the design allows for each degree of freedom to be associated with a single effect or hypothesis.

Source	Null Hypothesis	
A	$\mu_{A_1} = \mu_{A_2}$	1
B	$\mu_{B_1} = \mu_{B_2}$	1
C	$\mu_{C_1} = \mu_{C_2}$	1
AB	$\mu_{A_1B_1} + \mu_{A_2B_2} = \mu_{A_1B_2} + \mu_{A_2B_1}$	
AC	$\mu_{A_1C_1} + \mu_{A_2C_2} = \mu_{A_1C_2} + \mu_{A_2C_1}$	
BC	$\mu_{B_1C_1} + \mu_{B_2C_2} = \mu_{A_1C_2} + \mu_{A_2C_1}$	
ABC	$\mu_{A_1B_1C_1} + \mu_{A_1B_2C_2} + \mu_{A_2B_1C_2} + \mu_{A_2B_2C_1} =$ $\mu_{A_1B_1C_2} + \mu_{A_1B_2C_1} + \mu_{A_2B_1C_1} + \mu_{A_2B_2C_2}$	
SBS/ABC		24
Total		31

Degrees of freedom and sum of squares can be computed in the conventional way.
Thus

$$SS_S = \frac{(62)^2}{16} + \frac{(69)^2}{16} - \frac{(131)^2}{32} = 1.53$$

$$SS_T = \frac{(77)^2}{16} + \frac{(54)^2}{16} - \frac{(131)^2}{32} = 16.53$$

$$SS_{SxT} \text{ cells} = \frac{(29)^2}{8} + \frac{(33)^2}{8} + \frac{(48)^2}{8} + \frac{(21)^2}{8} - \frac{(131)^2}{32} = 48.09$$

$$SS_{ST} = SS_{SxT \text{ cells}} - SS_{sex} - SS_T = 48.09 - 1.53 - 16.53 = 30.03$$

Other two-way interactions would be calculated similarly. The three-way interaction SS would be calculated by subtracting all other sums of squares from the SS_{STC} cells.

The SS_{Total} would equal $\Sigma X^2 - \frac{(\Sigma X)^2}{32}$

The $SS_{SBS/ABC}$ would equal $SS_{Total} - SS_{ABC}$ cells

Because there is one df for each effect, it is possible to calculate sums of squares for each hypothesis (effect, df). The weights and calculations are as follows:

S	T	C		S	T	C	ST	SC	TC	STC	ST/C	ST/U
M	T	C	16	+1	+1	+1	+1	+1	+	+1	+	
		U	13	+1	+1	-1	+1	-	-	-1		+
	N	C	23	+1	-1	+1	-1	+	-	-1	-	
		U	10	+1	-1	-1	-1	-	+	+1		-
F	T	C	37	-1	+1	+1	-1	-	+	-1	-	
		U	11	-1	+1	-1	-1	+	-	+1		-
	N	C	13	-1	-1	+1	+1	-	-	+1	+	
		U	8	-1	-1	-1	+1	+	+	-1		+
Σ (W)(X)				62 69	77 54	89 42	50 81	58 73	71 60	50 81	29 60	21 21
$\Sigma(W)(\Sigma X)^2$		D D^2		7 49	23 529	47 2209	31 961	15 225	11 121	31 961	31 961	21 21
SS				1.53	16.53	69.03	30.03	7.03	3.78	30.03	60.06	SS Between 157.96

The summary for this analysis is as follows:

Source	df	SS	MS	F	
Sex	1	1.53	1.53		
Training	1	16.53	16.53	4.02	
Coord	1	69.03	69.03	16.77	< .01
ST	1	30.03	30.03	7.30	< .01
SC	1	7.03	7.03	1.71	
TC	1	3.78	3.78		
STC	1	30.03	30.03	7.30	< .01
(Between Cells)	(7)	(157.97)			
Within	24	98.75	4.11		
Total	31	256.72			

There is clearly a significant effect due to coordination, a significant interaction of sex by training, and a triple-order (ABC) interaction.[32]

Sex by training by coordination interaction. This pattern of results is an interpretive nightmare. The S x T interaction indicates that training had different effects on different sexes. However, this interpretation must be qualified by the significance of the three-way interaction which indicates that the S x T interaction is different at different levels of C.

There are at least three conventional ways to deal with significant interactions:

(1) plot dates with no further analysis;
(2) Repartition significant effects in order to simplify the hypotheses that are posed; or
(3) use multiple t-tests or some more conservative variation of multiple t's in order to test all means against all other.

Plotting significant effects: It is conventional to plot significant effects with the dependent variable on the

[32] Heal, op. cit., pp. 28-32.

vertical axis and the independent variable on the horizontal. The interpretation that would be made simply from plotting the effects would be that coordinated subjects did better than uncoordinated subjects; training produced improvement in females but, if anything, produced a decrement in males; and finally, the relative advantage of training for females held mainly for coordinated subjects and not for untrained subjects.

Repartitioning total mean square: If there is a ST interaction, then the two sexes must respond differently to different training procedures. A very straightforward statistical way of dealing with this interaction would be to ask about the two training effects for the two sexes separately. The comparison of simple effects (T/M and T/F) involves two degrees of freedom. In general, it will involve the number of degrees of freedom for the first factor times the number levels of the second factor.

These new variance estimators are derived from repartition of the main effect of the first (T) and the interaction of the first with the second (TS). Thus, in general

$$df_A + df_{AB} = df_{A/B_1} + df_{A/B_2} + \ldots + df_{A/B_b}$$

and

$$SS_A + SS_{AB} = SS_{A/B_1} + SS_{A/B_2} + \ldots + SS_{A/B_b}$$

The sums of squares for these simple effects are calculated as follows:

$$SS_{T/F} = \frac{(48)^2}{8} + \frac{(21)^2}{8} - \frac{(69)^2}{16} \doteq 45.5$$

or, if there is only one df, the weights can be used.

$$\frac{(48 - 21)^2}{(8)(2)} = \frac{(27)^2}{16} = 45.63$$

$$SS_{T/M} = \frac{(29)^2 + (33)^2}{(8)(18)} - \frac{(62)^2}{16} = 1.00$$

Thus, training produced a significant difference for females but not for males.

Note that 45.63 + 1.00 is equal to 16.53 + 30.03. Thus, the new partition uses the same overall variance as the original T and ST effects.

The effect of training within females is significant, $F(1,24) = 11.10$, whereas that for males is not, $F(1,24) < 1.00$. The interpretation placed on these simple effects is thus parallel to that resulting from a graph. Training was effective for females but not for males.

While these calculations demonstrate the use of 'simple effects,' the present situation calls for something more complicated. As noted above, the interpretation of the TSC interaction is that the TS interaction is different for different levels of C. To refine this interpretation we could repartition TS and TSC into TS/C and TS/U. This partition would indicate whether or not the TS interaction was significant for coordinated individuals, for uncoordinated individuals, or for both. However, this partition does not break down the TS interaction.

In complex cases like the present one, it is sometimes desirable to make both the TS and TSC re-partitions simultaneously. Thus, the T, TS, TC, and TSC effects from the original analysis can be re-partitioned into T vs. N/M/C, T vs. N/M/U, T vs. N/F/C, and T vs. N/F/U. The calculation of these simple effects is the same as the others.

$$SS_T \text{ vs. N/M/C} = \frac{(16)^2}{4} + \frac{(23)^2}{4} = \frac{(39)^2}{8} = 6.125$$

or using weights

$$SS_T \text{ vs. N/M/C} = \frac{(16 - 23)^2}{(2)(4)} = \frac{7^2}{8} = 6.125$$

Similarly:

$$SS_T \text{ vs. N/M/U} = \frac{(13 - 10)^2}{(2)(4)} = \frac{9}{8} = \underline{\underline{1.125}}$$

$$SS_T \text{ vs. N/F/C} = \frac{(37 - 13)^2}{(2)(4)} = \frac{(24)^2}{8} = \underline{\underline{72.00}}$$

$$SS_T \text{ vs. N/F/U} = \frac{(11 - 8)^2}{(2)(4)} = \frac{9}{} = \underline{\underline{1.125}}$$

Thus, these two interactions can be traced to the simple finding that Training helps coordinated females but not uncoordinated females or either kind of males.

Checking that the new partition is equivalent to the original partition,[33]

$$SS_{T/M/C} + SS_{T/M/U} + SS_{T/F/C} + SS_{T/F/U} = SS_T + SS_{TS} + SS_{TC} + SS_{TSC}$$

$$6.125 + 1.125 + 72.00 + 1.125 = 16.53 + 30.03 + 3.78 + 30.03 = 80.37$$

---

[33] Ibid., pp. 33-38.

Quasi-Experimental Designs

The Time-Series Experiment (7.0). The essence of the time-series design is the presence of a periodic measurement process on some group or individual and the introduction of an experimental change into this time series of measurement, the results of which are indicated by a discontinuity in the measurements recorded in the time series. It can be diagramed thus:

$$O_1\ O_2\ O_3\ O_4\ \ X\ \ O_5\ O_6\ O_7\ O_8$$

. . . .

The failure to control history is the most definite weakness of Design 7. That is, the rival hypothesis exists that not X but some more or less simultaneous event produced the shift. It is upon the plausibility of ruling out such extraneous stimuli that credence in the interpretation of this experiment in any given instance must rest. . . . In many situations in which Design 7 might be used, the experimenter could plausibly claim experimental isolation in the sense that he was aware of the possible rival events that might cause such a change and could plausibly discount the likelihood that they explained the effect.

Among other extraneous variables which might for convenience be put into history are the effects of weather and the effects of season. Experiments of this type are apt to extend over time periods that involve seasonal changes and, as in the studies of worker output, the seasonal fluctuations in illumination, weather, et cetera, may be confounded with the introduction of experimental change. Perhaps best also included under history, although in some sense akin to maturation, would be periodical shifts in the time series related to institutional customs of the group, such as the weekly work-cycles, pay-period cycles, examination periods, vacations, and student festivals. The observational series should be arranged to hold known cycles constant, or else be long enough to include several such cycles in their entirety.

. . . .

Many hypotheses invoking changes in instrumentation would lack a specific rationale for expecting the instrument error to occur on this particular occasion, as opposed to earlier ones. . . . In most instances, to preserve the interpretability of a time-series, it would be better to continue to use a somewhat antiquated device rather than to shift to a new instrument.

Regression effects are usually a negatively accelerated function of elapsed time and are therefore implausible as explanations of an effect at O_5 greater than the effects at O_2, O_3, and O_4. Selection as a source of main effects is ruled out in both this design and in Design 2, if the same specific persons are involved at all Os. . . .

Regarding external validity, it is clear that the experimental effect might well be specific to those populations subject to repeated testing. . . .

. . . .

It also seems imperative that the X be specified before examining the outcome of the time series. . . .

The prevalence of this design in the more successful sciences should give us some respect for it, yet we should remember that the facts of 'experimental isolation' and 'constant conditions' make it more interpretable for them than for us. It should also be remembered that, in their use of it, a single experiment is never conclusive. . . .[34]

Reliability (7.1). Reliability is consistency. If a particular test yields the same score every time it is administered, it is said to be reliable. There are two mathematically related ways to measure reliability. First, a reliability is the correlation between pairs of scores on the same individuals on two occasions.

For example, an index to the reliability or consistency of four subjects from one occasion to another could be obtained by correlation of their scores from the two occasions:

	0_1	0_2	Σ	0_1^2	$0_1 0_2$	0_2^2
S_1	2	3	5	4	6	9
S_2	4	6	10	16	24	36
S_3	1	2	3	1	2	4
S_4	0	3	3	0	0	9
ΣX	7	14	21	21	32	58
ΣX^2	21	58	143			

$$r_{0_1 0_2} = \frac{n\,\Sigma 0_1 0_2 - \Sigma 0_1\,\Sigma 0_2}{\sqrt{[\,n\,\Sigma 0_1^2 - (\Sigma 0_1)^2]\;[\,n\Sigma 0_2^2 - (\Sigma 0_2)^2\,]}}$$

[34]Campbell and Stanley, op. cit., pp. 37-42.

$$= \frac{(4)(32) - (7)(14)}{\sqrt{[\ (4)(21) - (7)(7)\ ]\ [\ (4)(58) - (14)(14)\ ]}}$$

$$= \frac{5}{\sqrt{35}} \doteq .845$$

Second, reliability can be seen as an index of the extent to which subjects interact with occasions in an analysis of variance sense. If subject X occasions interactions that are small relative to individual differences among subjects, then their scores are seen as reliable. Thus, in the preceding example:

$$SS_{SBS} = \frac{143}{2} - \frac{(21)^2}{2} = 16.37$$

$$SS_b = \frac{(7)^2 + (14)^2}{4} - \frac{(21)^2}{2} = 6.13$$

$$SS_{Tot} = 21 + 58 - \frac{(21)^2}{2} = 79.0 - 55.12 = 23.88$$

$$SS_{Oxsbs} = 23.88 - 16.37 - 6.13$$

$$r = \frac{MS_{sbs} - MS_{sbsx0}}{MS_{sbs} + MS_{sbsx0}} = \frac{\frac{16.37}{3} - \frac{1.38}{3}}{\frac{16.37}{3} + \frac{1.38}{3}} = \frac{14.99}{17.75} = .845$$

Thus, the correlation between two sets of scores is mathematically equivalent to the ratio of the MS between score pairs minus the MS for the Pairs x Occasions divided by the MS Pairs plus the MS_{PxO}. In other words, the reliability indexes the amount of variability between subjects after adjustments are made for inconsistencies within subjects from occasion 1 to occasion 2. It is clear that these 'occasions' can be raters, so that the reliability of raters can be estimated on alternate test forms, so that the reliability of alternate forms can be estimated. The analysis of variance approach to reliability has two features that are not offered by the correlational approach. First, several occasions can be used, and the reliability index per occasion or per n occasions can be calculated.

Suppose, in the second case, that there are two raters for each of the two occasions for each of ten subjects. The analysis of variance summary would be:

SBS	9
Raters	1
R x S	9
Occasions	1
0 x S	9
R x 0	1
R x 0 x S	9
	39

$$r_{Raters} = \frac{MS_{SBS} - MS_{RxS}}{MS_{SBS} + MS_{RxS}}$$

Given that one rater might be more lenient than another, are subjects seen to have consistent individual differences by two raters?

$$r_{Occasions} = \frac{MS_{SBS} - MS_{0xS}}{MS_{SBS} + MS_{0xS}}$$

Given that one occasion might give better performance than another, do subjects maintain consistent individual differences on the two occasions?

$$r_{Rx0} = \frac{MS_{Rx0} - MS_{Rx0xS}}{MS_{Rx0} + MS_{Rx0xS}}$$

Given that raters might change their sets on two occasions (Rx0), raters might vary in leniency (R), and one occasion might produce better performance than another (0), are these effects consistent?[35]

[35] Laird Heal, "Reliability Notes" (Nashville, Tennessee: George Peabody College, 1969), pp. 37-40. (Mimeographed.)

Chapter 2

ADVANCED RESEARCH DESIGNS AND METHODS FOR EDUCATIONAL ADMINISTRATORS

In this study, a major distinction was made between a practitioner and a basic researcher. The purpose of the basic researcher is to develop theory and create new knowledge. A practitioner applies research methodology and theory to the resolution of immediate and practical problems. While it is generally agreed that theory can suggest what variables are important in undertaking to solve an applied research problem, the development of theory, the initial design of research techniques, and verification of theory are not the province of most educational administrators or field practitioners of research.

Various types of research activities and their relationship to educational administration from project development to demonstration have been suggested as appropriate program components in the training of educational researchers. Table 2 summarizes three categories of research which contribute to both basic and applied problems. The practitioner and the basic researcher may use any or all of the three types of research described in Table 2.

The first chapter of the second volume of this study outlines program components for those planning to be practicing school administrators or consumers of research. Research designs and program components which follow in this chapter are presented for those planning to pursue a more in-depth study of basic educational research.

Table 2

Types of Research and Their Relationship to Educational Innovation*

	ANALYTIC	DESCRIPTIVE	EXPERIMENTAL
Purpose	To derive relationships within a deductive system	To describe existing conditions	To test casual relationships
Methods	Deductive, mathematical, historical philosophical, legal linguistic	Correlations, surveys, case studies, direct observation, cross cultural, growth studies	Comparison of experimental and non-experimental groups by systematically varying conditions
Relation to Innovation	Points out assumptions and possible consequences of proposed changes; useful in establishing criteria	Describes currently existing conditions so that they can be modified later	Shows the effects of a proposed innovation

*Gilbert Sax, Empirical Foundations of Educational Research (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), p. 36.

Quasi-Experimental Designs

The Equivalent Time-Samples Design (8.0). The most commonly used form of experimental design employs an equivalent sample or control group to provide a baseline against which to compare the effects of the experimental variable. In contrast, a recurrent form of one-group experimentation employs two equivalent samples of occasions, in one of which the experimental variable is present and in another of which it is absent. This design can be diagramed as follows (although a random rather than a regular alternation is intended):

$$X_1 0 \quad X_0 0 \quad X_1 0 \quad X_0 0$$

This design can be seen as a form of the time-series experiment with the repeated introduction of the experimental variable. The experiment is most useful where the effect of the experimental variable is anticipated to be of transient or reversible character. . . .

Most experiments employing this design have used relatively few repetitions of each experimental condition, but the type of extension of sampling theory represented by Brunswik (1956) calls attention to the need for large, representative, and equivalent random samplings of time periods.

As employed by Kerr, Design 8 seems altogether internally valid. History, the major weakness of the time-series experiment, is controlled by presenting X on numerous separate occasions, rendering extremely unlikely any rival explanation based on the coincidence of extraneous events. The other sources of invalidity are controlled by the same logic detailed for Design 7. With regard to external validity, generalization is obviously possible only to frequently tested populations. The reactive effect of arrangements, the awareness of experimentation, represents a particular vulnerability of this experiment. . . . Regarding the interaction between selection and X: there is, as usual, the limitation of the generalization of the demonstrated effects of X to the particular type of population involved.

This experimental design carries a hazard to external validity which will be found in all of those experiments in this paper in which multiple levels of X are presented to the same set of persons. This effect has been labeled 'multiple-X interference.' The effect of X_1, in the simplest situation in which it is being compared with X_0, can be generalized only to conditions of repetitious and spaced presentations of X_1. . . .

. . . .

This approach could be applied to a sampling of occasions for a single individual.[1]

---

[1]Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally and Company, 1968), pp. 43-45.

The Equivalent Materials Design (9.0). The design may be indicated in this fashion:

M_aX_10 M_bX_00 M_cX_10 M_dX_00 etc.

The Ms indicate specific materials, the sample M_a, M_c, etc., being, in sampling terms, equal to the sample M_b, M_d, etc. The importance of the sampling equivalence of the two sets of materials is perhaps better indicated if the design is diagramed in this fashion:

one person	Materials Sample A (0) X_0 0
or group	Materials Sample B (0) X_1 0

The 0s in parentheses indicate that in some designs a pretest will be used and in others not.

. . . .

Like Design 8, Design 9 has internal validity on all points, and in general for the same reasons. One may note, with regard to external validity, that the effects in Design 9, like those in all experiments involving repeated measures, may be quite specific to persons repeatedly measured. In learning experiments, the measures are so much a part of the experimental setting in the typical method used today (although not necessarily in Jost's method, in which the practices involved controlled numbers of readings of the lists) that this limitation on generalization becomes irrelevant. Reactive arrangements seem to be less certainly involved in Design 9 than in Design 8 because of the heterogeneity of the materials and the greater possibility that the subjects will not be aware that they are getting different treatments at different times for different items. . . .[2]

The Nonequivalent Control Group Design (10.0). One of the most widespread experimental designs in educational research involves an experimental group and a control group both given a pretest and a posttest, but in which the control group and the experimental group do not have pre-experimental sampling equivalence. Rather, the groups constitute naturally assembled collectives such as classrooms, as similar as availability permits but yet not so similar that one can dispense with the pretest. The assignment of X to one group or the other is assumed to be random and under the experimenter's control.

0 — X — — 0 —
0 0

. . . .

The hypothesis of an interaction between selection and maturation will occasionally be tenable even where the groups

---

[2]Ibid., pp. 46-47.

are identical in pretest scores. The commonest of these instances will be where one group has a higher rate of maturation or autonomous change than the other. Design 14 offers an extension of 10 which would tend to rule this out.

Regression provides the other major internal validity problem for Design 10. This hazard is avoidable, but one which is perhaps more frequently tripped over than avoided. . . .

. . . .

The threat of testing to external validity is as presented for Design 4. The interaction of selection and X reminds us that the effect of X may well be specific to respondents selected as the ones in our experiment have been. . . . The threat to external validity represented by reactive arrangements is present, but probably to a lesser degree than in most true experiments, such as Design 4.[3]

. . . .

Counterbalanced Designs (11.0). Under this heading come all of those designs in which experimental control is achieved or precision enhanced by entering all respondents (or settings) into all treatments. Such designs have been called 'rotation experiments,' 'counterbalanced designs,' 'cross-over designs,' and 'switch-over designs.' The Latin-square arrangement is typically employed in the counterbalancing. Such a Latin square is employed in Design 11, diagramed here as a quasi-experimental design, in which four experimental treatments are applied in a randomized manner in turn to four naturally assembled groups or even to four individuals:

	Time 1	Time 2	Time 3	Time 4
Group A	$X_1 0$	$X_2 0$	$X_3 0$	$X_4 0$
Group B	$X_2 0$	$X_4 0$	$X_1 0$	$X_3 0$
Group C	$X_3 0$	$X_1 0$	$X_4 0$	$X_2 0$
Group D	$X_4 0$	$X_3 0$	$X_2 0$	$X_1 0$

. . . The design contains three classifications (groups, occasions, and Xs or experimental treatments). Each classification is orthogonal to the other two in that each variate of each classification occurs equally often (once for a Latin square) with each variate of each of the other classifications. To begin with, it can be noted that each treatment (each X) occurs once and only once in each column and only once in each row. The same Latin square can be turned so that Xs become row or column heads, e.g.:

---

[3]Ibid., pp. 47-50.

	X_1	X_2	X_3	X_4
Group A	$t_1 0$	$t_2 0$	$t_3 0$	$t_4 0$
Group B	$t_3 0$	$t_1 0$	$t_4 0$	$t_2 0$
Group C	$t_2 0$	$t_4 0$	$t_1 0$	$t_3 0$
Group D	$t_4 0$	$t_3 0$	$t_2 0$	$t_1 0$

Sums of scores by Xs thus are comparable in having each time and each group represented in each. The differences in such sums could not be interpreted simply as artifacts of the initial group differences or of practice effects, history, et cetera. Similarly comparable are the sums of the rows for intrinsic group differences, and the sums of the columns of the first presentation for the differences in occasions. . . .

. . . There are systematic selection factors involved in the natural assemblage of the groups. These factors can be expected both to have main effects and to interact with history, maturation, practice effects, et cetera. . . . A second possible source of effects confounded with groups is that associated with specific sequences of treatments. . . .

Occasions are likely to produce a main effect due to repeated testing, maturation, practice, and cumulative carry-overs, or transfer. History is likewise apt to produce effects for occasions. The Latin-square arrangement, of course, keeps these main effects from contaminating the main effects of Xs. . . .

. . . .

Like all quasi-experiments, this one gains strength through the consistency of the internal replications of the experiment.[4]

<u>The Separate-Sample Pretest-Posttest Design (12.0)</u>. The design may be diagramed in this way:

```
R  0  (X)
R      X   0
```

In this paradigm, rows represent randomly equivalent subgroups, the parenthetical X standing for a presentation of X irrelevant to the argument. One sample is measured prior to the X, and an equivalent one subsequent to X. . . . While it has been called the 'simulated before-and-after design,' it is well to note its superiority over the ordinary before-and-after design (Design 2) through its control of both the main effect of testing and the interaction of testing with X. The main weakness of the design is its failure to control for history. . . .

[4]Ibid., pp. 50-52.

. . . Repeating Design 12 in different settings at different times, controls for history, in that if the same effect is repeatedly found, the likelihood of its being a product of coincidental historical events becomes less likely. . . . By replicating the effect under other settings, one can reduce the possibility that the observed effect is specific to the single population initially selected.

Maturation is unlikely to be invoked as a rival explanation, even in a public opinion survey study extending over months. . . .

Instrumentation represents a hazard in this design when employed in the sample survey setting. If the same interviewers are employed in the pretest and in the posttest, it usually happens that many were doing their first interviewing on the pretest and are more experienced, or perhaps more critical, on the posttest. If the interviewers are aware of the hypothesis, and whether or not the X has been delivered, then interviewer expectations may create differences. . . .

For pretests and posttests separated in time by several months, mortality can be a problem in Design 12. . . .

Perhaps for studies over long periods the pretest and posttest samples should be selected independently and at appropriately different times, although this, too, introduces a source of systematic bias resulting from possible changes in the residential pattern of the universe as a whole. . . .

It is characteristic of this design that it moves the laboratory into the field situation to which the researcher wishes to generalize, testing the effects of X in its natural setting.[5]

<u>The Separate-Sample Pretest-Posttest Control Group Design (13.0)</u>. It is expected that Design 12 will be used in those settings in which the X, if presented at all, must be presented to the group as a whole. If there are comparable (if not equivalent) groups from which X can be withheld, then a control group can be added to Design 12, creating Design 13:

```
R  0  (X)
R      X  0
- - - - - - -
R  0
R         0
```

As with Design 10, the weakness of Design 13 for internal validity comes from the possibility of mistaking for an effect of X a specific local trend in the experimental group which is, in fact, unrelated. By increasing the number of the social units involved (schools, cities, factories, ships, et cetera) and by assigning them in some number and with randomization to the experimental and control treatments, the one source of invalidity can be removed, and a true experiment, like Design 4

[5]Ibid., pp. 53-54.

except for avoiding the retesting of specific individuals, can be achieved. . . .

This design receives a perfect score for both internal and external validity.[6]

The Multiple-Time-Series Design (14.0). In studies of major administrative change by time-series data, the researcher would be wise to seek out a similar institution not undergoing the X, from which to collect a similar 'control' time series (ideally with X assigned randomly):

```
0 0 0 0X0 0 0 0
- - - - - - - -
0 0 0 0 0 0 0 0
```

This design contains within it (in the 0s bracketing the X) Design 10, the Nonequivalent Control Group Design, but gains in certainty of interpretation from the multiple measures plotted, as the experimental effect is in a sense twice demonstrated, once against the control and once against the pre-X values in its own series, as in Design 7. In addition, the selection-maturation interaction is controlled to the extent that, if the experimental group showed in general a greater rate of gain, it would show up in the pre-X 0s. Because maturation is controlled for both experimental and control series, by the logic discussed in the first presentation of the Time-Series Design 7, the difference in the selection of the groups operating in conjunction with maturation, instrumentation, or regression, can hardly account for an apparent effect. An interaction of the selection difference with history remains, however, a possibility.

. . . .

In general, this is an excellent quasi-experimental design, perhaps the best of the more feasible designs. . . .[7]

The Recurrent Institutional Cycle Design: A 'Patched-Up' Design (15.0). Design 15 illustrates a strategy for field research in which one starts out with an inadequate design and then adds specific features to control for one or another of the recurrent sources of invalidity. The result is often an inelegant accumulation of precautionary checks, which lacks the intrinsic symmetry of the 'true' experimental designs, but nonetheless approaches experimentation. As a part of this strategy, the experimenter must be alert to the rival interpretations (other than an effect of X which the design leaves open) and must look for analyses of the data, or feasible extensions of the data, which will rule these out. Another feature often characteristic of such designs is that the effect of X is demonstrated in several different manners. This is obviously

[6]Ibid., p. 55.

[7]Ibid., pp. 55, 57.

an important feature where each specific comparison would be equivocal by itself.

The specific 'patched-up' design under discussion is limited to a narrow set of questions and settings, and opportunistically exploits features of these settings. . . .

. . . .

In idealized form this design is as follows:

Class A	X	0_1		
Class B		0_2	X	0_3

This design combines the 'longitudinal' and 'cross-sectional' approaches commonly employed in developmental research. In this it is assumed that the scheduling is such that at one and the same time a group which has been exposed to X and a group which is just about to be exposed to it can be measured; this comparison between 0_1 and 0_2 thus corresponds to the Static-Group Comparison, Design 3. Remeasuring the personnel of Class B one cycle later provides the One-Group Pretest-Posttest segment, Design 2. The cross-sectional comparison of $0_1 > 0_2$ provides differences which could not be explained by the effects of history or a test-retest effect. The differences obtained could, however, be due to differences in recruitment from year to year. . . .

. . . .

The design as represented through measurements 0_1 to 0_5 uniformly fails to control for maturation. The seriousness of this limitation will vary depending upon the subject material under investigation.

. . . In such a situation a control for maturation seems very essential. For this reason 0_6 and 0_7 have been added to the design, to provide a cross-sectional test of a general maturation hypothesis made on the occasion of the second testing period. This would involve testing two groups of persons from the general population who differ only in age and whose ages were picked to coincide with those of Class B and Class C at the time of testing. To confirm the hypothesis of an effect of X the groups 0_6 and 0_7 should prove to be equal, or at least to show less discrepancy than do the comparisons spanning exposure to X. . . .

Another cross-sectional approach to the control of maturation may be available if there is heterogeneity in age (or years away from home, etc.) within the population entering the institutional cycle. . . .

. . . .

In the diagrams of Design 15 as presented, it is assumed that it would be feasible to present the posttest for one group at the same chronological time as the pretest for another. This is not always the case in situations where one might want to use this design. The following is probably a more accurate portrayal of the typical opportunity in the school situation:

Class A	X O_1		
Class B_1		RO_2 X O_3	
Class B_2		R X O_4	
Class C			O_5 X

. . . .

If the cross-sectional and longitudinal comparisons indicate comparable effects of X, this could not be explained away as an interaction between maturation and the selection differences between the classes. Since the X is so complex, the investigation is apt to be made for practical reasons rather than theoretical purposes, and for these practical purposes, it is probably to this one institution that one wants to generalize in this case.[8]

Regression-Discontinuity Analysis (16.0). This is a design developed in a situation in which ex post facto designs were previously used. While very limited in range of possible applications, its presentation here seems justified as an illustration of the desirability of exploring in each specific situation all of the implications of a causal hypothesis, seeking novel out-croppings where the hypothesis might be exposed to test. . . .

. . . .

Some of the tests of significance discussed for Design 7 are relevant here. Note that the hypothesis is clearly one of intercept difference rather than slope, and that the location of the step in the regression line must be right at the X point, no 'lags' or 'spreads' being consistent with the hypothesis. Thus parametric and nonparametric tests avoiding assumptions of linearity are usually more plausible for such regression data than for time series. This might make a t test for the difference between the two linearly extrapolated points appropriate. Perhaps the most efficient test would be a covariance analysis, in which the award-decision scores would be the covariate of later achievement, and award and no-award would be the treatment.

Is such a design likely to be used? It certainly applies to a recurrent situation in which claims for the efficacy of X abound. . . .

. . . .

Because of synchrony of experimental and control group, history and maturation seem controlled. Testing as a main effect is controlled in that both the experimental and control groups have received it. Instrumentation errors might be a problem if the follow-up O were done under the auspices which made the award, in that gratitude for the

---

[8] Ibid., pp. 57-61.

award and resentment for not receiving the award might lead to differing expressions of attitude, differing degrees of exaggeration of one's own success in life, et cetera. This weakness would also be present in the tie-splitting true experiment. It could be controlled by having the follow-ups done by a separate agency. It is believed, following the arguments above, that both regression and selection are controlled as far as their possible spurious contributions to inference are concerned, even though selection is biased and regression present; both have been controlled through representing them in detail, not through equation. Mortality would be a problem if the awarding agency conducted the follow-up measure, in that award recipients, alumni, et cetera, would probably cooperate much more readily than non-winners. Note how the usually desirable wish of the researcher to achieve complete representation of the selected sample may be misleading here. . . . For both, the selection-maturation interaction threat to internal validity is controlled. For the quasi-experiment, it is controlled in that this interaction could not actually explain a distinct discontinuity in the regression line at X. The external validity threat of a testing-X interaction is controlled to the extent that the basic measurements used in the award decision are a part of the universe to which one wants to generalize.

Both the tie-breaking true experiment and the regression-discontinuity analysis are particularly subject to the external-validity limitation of selection-X interaction in that the effect has been demonstrated only for a very narrow band of talent, i.e., only for those at the cutting score.[9]

Descriptive Research Designs

The Survey. The sample survey is a technique involving larger numbers of persons than the case study. It attempts to describe population characteristics by selecting an unbiased sample and generalizing the results of the sample to the population from which it is drawn. Often, the population is sampled by using questionnaires, although interviews or tests may also be used.

Problems of sampling and questionnaire development are two essential components of the sample survey. The student should remember that the survey is a type of descriptive research in which information is obtained from a sample of respondents in order to test hypotheses concerning the status of some educational problem. An example might be an administrator's interest in knowing the attitudes of school board members throughout the United States to collective negotiations.

---

[9]Ibid., pp. 61-63.

Because of the difficulty of studying an entire population of individuals, objects, or methods, the educational researcher usually selects samples by methods that attempt to ensure unbiased, suitably close approximations to the relevant characteristics of the sampled population.

Ideally, the population to which one plans to generalize research findings determines the selection of the sample. In actual practice, however, it is usually necessary to limit the scope of a population so that it may be sampled.

The process of sampling consists of three phases: (1) defining the population; (2) drawing a sample from the population; and (3) estimating the mean, standard deviation, proportion, or other parameter values from knowledge obtained from the sample (statistical inference).

A sample is described as being unbiased when elements are drawn in some random manner. Unbiased samples approximate parameters as additional cases are drawn; biased samples consistently overestimate or underestimate parameters.

In simple random sampling, every element in the population has an equal probability of being selected. The use of a table of random numbers is an efficient and easy way of drawing a simple random sample.

With simple random sampling, the standard error of a mean is a ratio of the standard deviation to the square root of the number of cases minus 1. Thus one can increase the precision of the parameter estimate by reducing the variability of the sample and/or by increasing the number of cases.

Stratified random sampling involves: (1) dividing a population into strata; (2) sampling randomly within each stratum; and (3) estimating the value of the parameter. The major advantage to stratification is that sampling error arises only within each stratum and not between different strata. In addition, stratification allows the use of different methods of drawing samples within each stratum, which may help to reduce costs. Methods are available for estimating parameters for proportional and disproportional stratified sampling. Disproportional sampling requires a weighting procedure for estimating parameters; weighting is not necessary for proportional stratified sampling.

Systematic sampling is the selection of every nth case from a population listing or roster. The principal advantage of systematic sampling over simple random sampling is one of convenience. The first number should be drawn randomly. The major disadvantage to systematic sampling is that it may be seriously in error if elements in the population are repeated at constant intervals. In addition, standard error formulas for systematic samples have not been satisfactorily developed.

In area or cluster sampling major areas or clusters are selected first and individuals second. The major purpose of cluster sampling is to reduce costs per element sampled, although this advantage is somewhat offset by the usual increase in error. To reduce error, clusters would have to be selected in such a way that differences between clusters are small in comparison to the variability within clusters.

The size of a sample is determined by a number of different factors: accuracy, cost, and homogeneity.[10]

The Interview Technique. As a research method, the interview is more than an exchange of small talk. It represents a direct attempt by the researcher to obtain reliable and valid measures in the form of verbal responses from one or more respondents. The data obtained from interviews represent attempts to confirm or reject hypotheses. As such, they are a part of a research data collection procedure, and not ends in themselves.

A number of advantages have been claimed for the interview over other research techniques. First, the interview is highly flexible and is applicable to many different types of problems. It is flexible in the sense that the interviewer may change his mode of questioning if the occasion demands. . . .

. . . .

In addition to being flexible, the interview allows the investigator to observe both what the respondent has to say and the way in which he says it. If the interview is structured, or standardized, it is similar to the administration of individual intelligence tests. How the subject responds may be as important as the content of his responses.

A third advantage of the interview is its usefulness in collecting personal information, attitudes, perceptions, or beliefs, and in probing for additional information if needed. . . .

A fourth advantage of the interview concerns motivation. Almost all interviews begin with an attempt to develop rapport between the interviewer and the respondent. . . .

. . . The flexibility of the interview is a distinct advantage, but it also generates its own peculiar difficulties, especially where the interview is unstructured. . . .

A second problem concerns the training of the interviewer himself. Personal values, beliefs, and biases of observers may influence the outcomes of investigations. To compensate for these subjective factors, interviewers must be trained and evaluated. This training will add substantially to the costs of data collection.

A third difficulty is that the interview has often been used inappropriately where other more suitable techniques should have been employed. . . .

. . . .

The literature on interviewing contains many 'dos' and 'don'ts' which are either vague or unsubstantiated by research findings. However, Kahn and Cannell have suggested some of

---

[10]Gilbert Sax, Empirical Foundations of Educational Research (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968), pp. 149-50, 292.

the potential sources of error which may occur in any interview: errors in asking questions, errors in probing, errors in motivating, and errors in recording responses.[11]

The Questionnaire Technique. In many ways the interview and the questionnaire are similar. Both attempt to elicit the feelings, beliefs, experiences, or activities of respondents. They also may both involve formats which can be relatively structured or unstructured as the situation demands. . . .

. . . .

The major advantage of the questionnaire over the interview is one of economy. Because many questionnaires are sent through the mail, the expense and time involved in training interviewers and sending them personally to interview each respondent are diminished. Also, the questionnaire can be sent almost anywhere, a condition which is usually impractical where an interview is required or needed.

The use of the mails in sending out questionnaires means that more persons can be contacted than would be possible with interviews. . . .

Another advantage claimed for the questionnaire is that each respondent receives the same set of questions phrased in exactly the same way, in much the same way that the questions on standardized tests are posed. Questionnaires are thus supposed to yield data which are more comparable than information obtained by means of the interview. The validity of this argument depends, in part, on whether the questions are structured or unstructured. . . .

. . . .

There are a number of serious disadvantages in using questionnaires to obtain data. In the first place, the motivation of the respondent ordinarily cannot be checked, while in an interview situation rapport can be established. Without having some indication of the motivational level of respondents as they answer questions, the validity of their responses is difficult to judge.

A second disadvantage of the questionnaire, and especially of the mailed variety, is that its use assumes that respondents can both read and write. While this is probably not too serious a problem in many studies, it is a limitation if the population to be sampled contains illiterates or persons who are literate only in a foreign language.

A third disadvantage of the questionnaire is concerned with the problem of sampling. The difficulty usually lies not in choosing the sample, but in getting those persons selected for the sample to return their questionnaires. Because each questionnaire which is not returned may result in biased

---

[11]Ibid., pp. 201-04.

sampling, every effort should be made to obtain a 100 percent return. The usual expectation for mailed questionnaires is that fewer than half of the respondents may reply the first time they receive a questionnaire, the percentage of returns depending upon such factors as the length of the questionnaire, the reputation of the sponsoring agency, the complexity of the questions asked, the relative importance of the study as judged by the potential respondent, the extent to which the respondent believes that his responses are important, and the quality and design of the questionnaire itself.

. . . .

A questionnaire, no matter how well designed, does not represent an end in itself. It is a means of gathering information for specific purposes. . . . The decision to use a questionnaire (or any other instrument or method) should develop out of the investigator's hypotheses, which in turn should be justified by the criteria for the selection of a research problem. . . .

Questionnaires are tempting because of their apparent simplicity. After all, the student argues, all that is needed is to send a questionnaire to a large number of persons to determine their attitude toward some issue in education. However, the value of the proposed research is dependent upon the potential contribution of the study, which in turn is dependent upon the extent to which the study adds to or tests some aspect of educational theory or practice.

If the topic for investigation is related to a body of theory or can be justified on some other grounds, then the investigator faces the problem of deciding what type of information he wishes to collect. Because there are an infinite number of questions which could be asked on a questionnaire, the responsibility for limiting the questions rests on the researcher himself. One step which may be useful is the preparation of a work table containing a list of questions which may have some potential relevance to the hypotheses under investigation. For each of these questions, the investigator should indicate his justifications for collecting that type of information. . . .

. . . .

Following are descriptions of four types of measurement scales which can be used on questionnaires or other tests or inventories. These four types of scales are nominal, ordinal, interval, and ratio scales. . . .

The nominal scale involves the fewest assumptions. Names or labels for persons, objects, activities, or beliefs form this type of scale. Its purpose is simply to identify or categorize. . . .

The ordinal scale has all of the characteristics of the nominal scale, but in addition it allows us to place objects in rank order (highest to lowest, most to least, or best to worst, for example). . . .

. . . .

The interval scale can be ranked, and the differences between successive ranks will be equal. A thermometer is an example of interval measurements. Because intervals are equal, addition and subtraction are mathematically legitimate operations. . . .

. . . .

The ratio scale has not only equal intervals but also an absolute zero (i.e., no negatives). . . .

With ratio scales, all mathematical operations such as addition, multiplication, subtraction, and division are possible.

. . . .

The arrangement of items on a questionnaire depends, in part, on the type of attitude scale used. On a social distance scale, the seven categories of social distance (from 'To close kinship by marriage' to 'Would exclude from my country') are fixed by order by favorableness. . . .

. . . .

Campbell has distinguished four types of items: (1) nondisguised-structured, such as the Thurstone and Likert scales, in which the respondent is given accurate information about the purpose of the questionnaire but is restricted in his responses by the investigator; (2) nondisguised-nonstructured, as found on the free-response type of questionnaire, where the subject has unlimited possibilities to respond in any manner he sees fit; (3) disguised-nonstructured, which can be found on many projective techniques; and (4) disguised-structured, which restrict the responses of the subject and at the same time hide the real purpose of the questionnaire.

Item format can be considered from another point of view. For example, on achievement tests, it is customary to refer to items as true-false, multiple-choice, completion, rearrangement, or essay. The advantages and disadvantages of each of these types of items can also apply to questionnaires.

The true-false item is essentially a two-alternative statement which is scored as correct or incorrect. . . .

The Likert scale is essentially a multiple-choice test having three or more alternatives for each item. The subject is instructed to select the one option which best describes his opinion or belief. On a three-option questionnaire, the subject may either agree, disagree, or state that he is uncertain as to his beliefs.

On completion items, the subject is provided with an incomplete statement and is requested to write in a phrase or clause which will complete the sentence.

. . . .

The literature on item construction for questionnaires contains a large number of rules, or 'dos' and 'don'ts.' While many of these suggestions may be quite useful, they should be related to the purpose of the investigation. . . .

. . . .

Before a final form of the questionnaire is constructed, it is of advantage to conduct a pilot study to determine if the items are yielding the kind of information that is needed. . . .

If questionnaires are administered to an intact group (such as students in classes), the investigator has the opportunity to motivate respondents and to answer questions that may arise. However, because many questionnaires are sent through the mail, it is usually necessary to motivate respondents to fill out the questionnaire and to return it within a reasonable period of time. Unless respondents believe that the questionnaire is of value, that it is sponsored by a recognized organization, that their personal attention to the questionnaire does count and is important, and that not too much time will be required, the questionnaire is likely to be thrown into the nearest waste basket.

. . . .

The reliability of a set of measurements obtained from questionnaires may be determined by using coefficients of stability, equivalence, or internal consistency (homogeneity). . . .[12]

The Q Technique. The Q technique is another method of obtaining information about respondents. Q methodology begins with the development of items to be sorted into piles according to some method approximating a symmetrical distribution. Each pile is given a number ranging from zero to 10 (if 11 piles are being used) and the subject is instructed to place each card into a forced-choice distribution with more of the items or cards placed in the center of the distribution than at the ends. Correlations are then run between persons. This procedure makes it possible to study a few individuals at a time.

Although a number of studies have established the value of Q methodology, it has limitations. Forcing choices into a symmetrical distribution has been criticized, as has the sorting procedure.[13]

The Semantic Differential. The semantic differential was designed to measure the connotative meaning of concepts. Typically, developing a semantic differential involves establishing a number of bipolar adjectives which can be rated on a continuum. These scores are usually factor analyzed and a 'semantic space' defined.

The semantic differential allows the investigator not only to examine the connotative meaning of concepts, but to compare the profiles made by two or more persons. It is also possible

[12] Ibid., pp. 214-30.

[13] Ibid., pp. 281-82.

to compare the profiles of the same person on two or more concepts.

Although there are decided advantages in using the semantic differential, its limitations should be kept in mind. Being essentially a bipolar rating scale, it has all of the disadvantages that this type of scale is known to possess (such as ease of faking and of acquiescing), as well as its advantages (simplicity in administration and scoring).[14]

---

[14]Ibid., p. 282.

Chapter 3

WRITING AND INTERPRETING RESEARCH

During the present decade increased emphasis, generated by government programs and so-called educational innovations, has created an awareness of the need for better educational research.

One of the marks of a professional educator is a concern for good and properly reported educational research. The leading journal formats for reporting the writing and interpretation of research and experimentation all too often have been inadequate.

The writing and interpreting of most educational research activities is done to describe "what was" for the purpose of determining and describing historical development. In addition, comparisons or contrasts are reported descriptively for the purpose of interpreting "what is" or to describe the results of an experimental program in an attempt to answer "what will be." The research reporting usually involves schools or school children through action involving the practitioner on the job.

Prerequisite to any research report is the method comprising the problem definition. In Figure 7 the problem definition is outlined diagramatically and an approach to problem solving is provided in the accompanying material.

An approach to the writing and interpretation for the solution of a problem as described in Figure 7 may be reported using the outline which follows.

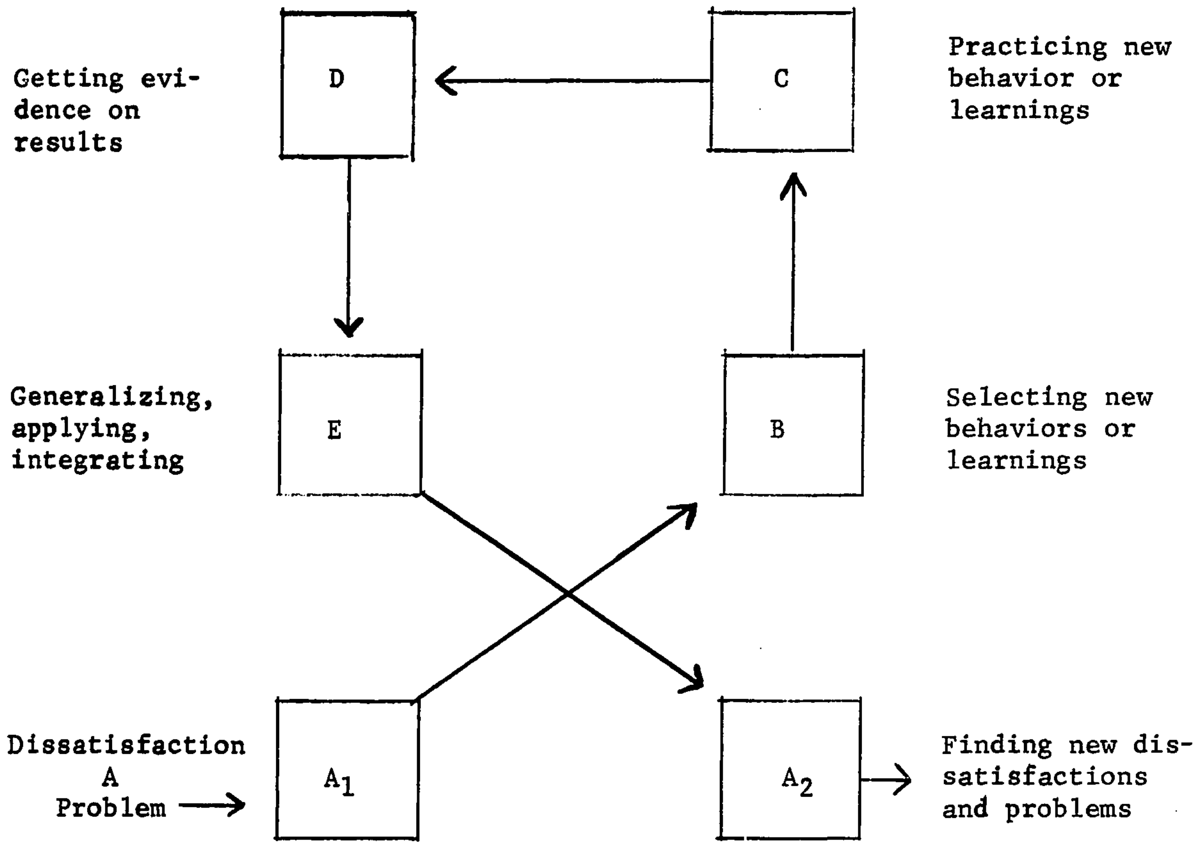

Figure 7

Schema for Problem Definition*

*Matthew B. Miles, Learning to Work in Groups (New York: Teachers College Bureau of Publications, Columbia University, 1959), p. 38.

Statement of the Problem

a. The problem should be stated in terms intelligible to someone who is generally sophisticated but who is relatively uninformed in the area of the problem.
b. Delimit the specific area to be investigated.
c. Formulate the questions to be raised and tested.
d. Specify the significance of the study:
   1. That is, indicate how this research project will refine, revise, or extend existing knowledge. These refinements, revisions, or extensions may have either substantive or methodological significance.
   2. Since all studies have two potential audiences (practitioners and professional peers) statements relating the research to both groups are in order.

Statement of the Purpose

One paragraph should be written indicating the purpose of the study, e.g., "The purpose of this paper (research, etc.) is . . . ."

Solution Design

a. What procedures (model) are needed to answer each question? What model, new model, or modified model is appropriate? Construct a model such as modified Delphi, or use PERT, CPM (or a modified version of these), or employ a cost/effectiveness approach to evaluate programs, or apply computer graphics, or utilize MBO to answer each question.
b. Present a thorough description of the selected model.
c. What type of data are needed? Will a sample set of real data or a simulated set of data be sufficient to use in the model?

Solution

a. What are the results of applying the data to the model?
b. List the alternatives that are produced (if appropriate, e.g., cost/effectiveness, PERT, CPM, Delphi, MBO).

Implement the Solution

Discuss the implementation of the solution and the advantages and disadvantages (pitfalls) of these solutions. Relate the advantages and disadvantages to the structure of the selected model.[1]

The writing and interpreting of research depends to a great extent on the experience, work style, and training of the investigator. In any research activity it is important to recognize that research is a systematic approach, resulting in an accurate investigation. The gathering of data from primary sources must be logically and objectively pursued. The qualitative organization of data must be carefully recorded and reported. Training programs for those preparing to write and interpret must adhere to the rigorous standards required in all research activities.

[1]C. Kenneth Tanner, "An Approach to Problem Solving" (Knoxville, Tennessee: University of Tennessee, 1971). (Mimeographed.)

Chapter 4

A SUGGESTED PROGRAM FOR EDUCATIONAL RESEARCH AS A MAJOR FIELD OF STUDY

Educational research, like research in the natural sciences, must deal with the practical and the theoretical. Educators, like scientists, depend upon research to provide sufficient information prerequisite to responsible action.

Throughout this study a distinction was made between a practitioner and a basic researcher. Close interaction between basic educational research and applied educational research is as much a necessity as it is in the natural sciences.

The training program for one planning to become a natural scientist involves a total approach, and includes, as a major part of the program, an emphasis on research. The training of educational researchers is analogous to that of natural scientists. For that reason, a suggested program for those planning educational research as a major field of study is presented at this point.

Program for the DOCTOR OF EDUCATION Degree (Major in Educational Research)

Area		Qtr. Hrs.
Statistics	Intro. to Statistical Methods	3
	Intermed. Statistical Methods	3
	Nonparametric Statistics	3
	Multivariate Statistics	3
	Intro. to Bayesian Statistics	3

Area		Qtr. Hrs.
Evaluation	Measurement and Evaluation	3
	Adv. Measurement and Evaluation	3
	Surveys and Rating Scales	3
	Theory of Measurement	3
Research Methods	Research Methods and Techniques	3
	Research Design	3
	Adv. Research Design	3
	Survey Research	3
	Intro. to Educational Planning	3
	Indiv. Study in Ed. Research	6
	Writing and Interpreting Research	3
	Computer Programming - FORTRAN	3
	Federal Programs & Grants	3
	Adv. Research & Analysis	3
Other Courses in:	Social Foundations of Education	3
	Curriculum or Media	3
Two minor fields and/or cognate work in Psychology, Sociology, Mathematics, etc.		48
Dissertation		24
		138

BIBLIOGRAPHY

BIBLIOGRAPHY

Volume II

Bruning, James L. and B. L. Kintz. Computational Handbook of Statistics. Glenview, Illinois: Scott, Foresman and Company, 1968.

Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Company, 1968.

Good, Carter V. Personal interview. November 30, 1971.

Heal, Laird. "ANOVA Lecture Notes." Nashville, Tennessee: George Peabody College, 1969. (Mimeographed.)

________. "Reliability Notes." Nashville, Tennessee: George Peabody College, 1969. (Mimeographed.)

________. "Covariance Notes." Nashville, Tennessee: George Peabody College, 1969. (Mimeographed.)

Hogge, James H. and Judith H. Picklesimer. "Peabody Statistical Library User's Manual." Nashville, Tennessee: George Peabody College, 1969. (Mimeographed.)

Miles, Matthew B. Learning to Work in Groups. New York: Teachers College Bureau of Publications, Columbia University, 1959.

Purdy, Ralph. Personal interview. December 1, 1971.

Sax, Gilbert. Empirical Foundations of Educational Research. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.

Tanner, C. Kenneth. "An Approach to Problem Solving." Knoxville, Tennessee: University of Tennessee, 1971. (Mimeographed.)

APPENDIXES

Volume II

APPENDIX A

SUGGESTED BIBLIOGRAPHY

RESEARCH DESIGN AND ANALYSIS OF RESEARCH PAPERS

American Association of School Administrators. Administrative Problems in Search of Researchers. Washington, D. C., 1966.

Barnes, Fred P. Research for the Practitioner in Education. Washington, D. C.: Department of Elementary School Principals, National Education Association, 1964.

Barnes, John S. Educational Research for Classroom Teachers. New York: Putnam, 1960.

Barr, Arvil Sylvester, Robert A. Davis, and Palmer O. Johnson. Educational Research and Appraisal. Philadelphia: Lippincott, 1953.

Bellack, Arno A. (ed.). Theory and Research in Teaching. New York: Bureau of Publications, Teachers College, Columbus University, 1963.

Best, John W. Research in Education. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1959.

Borg, Walter R. Educational Research, An Introduction. New York: D. McKay Company, 1963.

Budd, William C. and Sam P. Kelly. Educational Research by Practitioners: An Elementary Casebook. New York: Harper and Row, 1970.

________. Complete Guide and Index to ERIC, Reports thru December, 1969. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Conant, James Bryant. Trial and Error in the Improvement of Education. Washington, D. C.: Association for Supervision and Curriculum Development, 1961.

Culbertson, Jack A. and Stephen P. Hencley. Educational Research: New Perspectives. Danville, Illinois: Interstate Printers and Publishers, 1963.

________. Current Research on Instruction. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1969.

Curriculum Research Institute, Ninth San Francisco, 1963 and 1964. Theories of Instruction; Papers. Washington, D. C.: Association for Supervision and Curriculum Development, 1965.

Darr, George A. Characteristics of the Volunteer and the Nonvolunteer for Research Studies at East Tennessee State University. Johnson City, Tennessee: East Tennessee State University, 1967.

________. A Directory of Education Research Agencies and Studies. Bloomington, Indiana: Phi Delta Kappa, 1957.

Foskett, D. J. How to Find Out: Educational Research. Oxford, New York: Pergamon Press, 1965.

Fund for the Advancement of Education. Decade of Experiment: 1951-61. New York: The Fund for the Advancement of Education, 1961.

Gage, Nathaniel Lees. Handbook of Research on Teaching; a Project of the American Educational Research Association. Chicago: Rand McNally, 1963.

Galfo, Armand Jones and Earl Miller. Interpreting Educational Research. Dubuque, Iowa: W. C. Brown Company, 1965.

Gephart, William J. and Robert Ingle. Educational Research; Selected Readings. Columbus, Ohio: C. F. Merrill, 1969.

Good, Carter Victor. Essentials of Educational Research; Methodology and Design. New York: Appleton-Century-Crofts, 1966.

________. Introduction to Educational Research; Methodology of Design in the Behavioral and Social Sciences. New York: Appleton-Century-Crofts, 1963.

________, A. S. Barr, and Douglas E. Scates. The Methodology of Educational Research. New York: Appleton-Century-Crofts, 1941.

Grobman, Hulda Gross. Evaluation Activities of Curriculum Projects: a Starting Point. Chicago: Rand McNally, 1968.

Hayman, John L. Research in Education. Columbus, Ohio: C. E. Merrill Publishing Company, 1968.

Herman, Wayne L. Current Research in Elementary School Social Studies. New York: The Macmillan Company, 1969.

Hillway, Tyrus. Handbook of Educational Research; a Guide to Methods and Materials. Boston: Houghton Mifflin Company, 1969.

Hunnicutt, Clarence William and William J. Iverson (ed.). Research in the Three R's. New York: Harper and Row, 1958.

International Institute for Educational Planning. Educational Planning: an Inventory of Major Research Needs. Paris: International Inst'tute for Educational Planning, 1965.

Kelley, Clarice Y. Where It's Happening; a Selective Guide to Continuing Programs. Garden City, New York: Doubleday, 1968.

Krumboltz, John D. Learning and the Educational Process; Selected Papers from the Research Conference on Learning and the Educational Process, Stanford University, June 22-July 31, 1964. Chicago: Rand McNally, 1965.

Loughary, John William, John F. Cogswell, and others. Man-Machine Systems. New York: Harper and Row, 1966.

Maccia, Elizabeth Stainer. The Conceptions of Model in Educational Theorizing. Columbus: Bureau of Educational Research and Service, 1962.

Martin, William Ted and Dan C. Pinck. Curriculum Improvement and Innovation; a Partnership of Students, School Teachers, and Research Scholars. Cambridge, Massachusetts: R. Bentley, 1966.

Michigan State University, East Lansing, Office of Institutional Research. Evaluation in Higher Education. Boston: Houghton Mifflin Company, 1961.

Mouly, George J. The Science of Educational Research. New York: American Book Company, 1963.

National Research Council, Advisory Board on Education. Psychological Research in Education; Report of a Conference sponsored by the Advisory Board on Education, Easton, Maryland, April 24-26, 1958. Washington, D. C.: National Academy of Sciences, National Research Council, 1958.

Owens, Robert G. Organizational Behavior in Schools. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

________. Penn State Review of Educational Research, 1949. Pennsylvania: State University Pennsylvania Division of Educational Research, Department of Education, School of Education, 1949.

Phi Delta Kappan Symposium on Educational Research. Annual Phi Delta Kappa Symposium on Educational Research Proceedings. Bloomington, Indiana: Phi Delta Kappa, 1959.

Sax, Gilbert. Empirical Foundations of Educational Research. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.

Structure of Knowledge and the Nature of Inquiry Workshop, Williamette University, 1964. The Structure of Knowledge and the Nature of Inquiry; a Report. Salem: Division of Education Development, State Department of Education, 1965.

Taba, Hilda and Elizabeth Noel. Action Research, a Case Study. Washington, D. C.: Association for Supervision and Curriculum Development, 1957.

Thorndike, Robert Ladd. The Concepts of Over- and Underachievement. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Tolbert, E. L. Research for Teachers and Counselors; Reading Research, Carrying out Studies, and Employing the Approach in Daily Work in Graduate Courses and on the Job for Courses in Education, Guidance and Counseling. Minneapolis, Minnesota: Burgess Publishing Company, 1967.

Travers, Robert Morris. An Introduction to Educational Research. New York: The Macmillan Company, 1969.

U. S. Educational Research Information Center. Office of Educational Research Reports, 1956-65, ED002747-ED003960. Washington, D. C.: U. S. Government Printing Office.

U. S. Office of Education Bureau of Research. Report on Cooperative Research to Improve the Nation's Schools, 1965-1966. Washington, D. C.: U. S. Government Printing Office.

U. S. Office of Education. Cooperative Research Project, Bulletin. Washington, D. C.: U. S. Office of Education, 1956-.

U. S. Office of Education, Division of Higher Education. Special Reports, Clearinghouse of Studies on Higher Education. Washington, D. C.: U. S. Department of Health, Education and Welfare, Office of Education, Circular No. 563.

U. S. Office of Education. Research Projects of the Cooperative Research Program; a Seven-year Summary, July 1, 1956-June 30, 1963. Washington, D. C.: U. S. Office of Education, 1964.

U. S. Panel on Educational Research and Development. Innovation and Experiment in Education; a Progress Report to the U. S. Commissioner of Education, the Director of the National Science Foundation and the Special Assistant to the President for Science and Technology. Washington, D. C.: U. S. Government Printing Office, 1964.

Van Dalen, Deobold B. Understanding Educational Research. New York: McGraw-Hill Book Company, 1966.

Wiersma, William. <u>Research Methods in Education; an Introduction</u>. Philadelphia: Lippincott, 1969.

Wise, John Edward and others. <u>Methods of Research in Education</u>. Boston: D. C. Heath and Sons, 1967.

APPENDIX B

SUGGESTED BIBLIOGRAPHY

STATISTICAL TECHNIQUES

Alder, Henry L. and Edward B. Roessler. Introduction to Probability and Statistics. San Francisco: W. H. Freeman, 1964.

American Statistical Association. Committee on Social Statistics. Philadelphia: University of Pennsylvania Press, 1938.

Armore, Sidney J. Introduction to Statistical Analysis and Inference for Psychology and Education. New York: John Wiley and Sons, 1966.

Bernstein, Allen L. A Handbook of Statistics Solutions for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1964.

Blalock, Hubert M. Social Statistics. New York: McGraw-Hill Book Company, 1960.

Blommers, Paul J. and E. F. Lindquist. Elementary Statistical Methods in Psychology and Education. Boston: Houghton Mifflin Company, 1960.

Bowen, Earl K. Statistics with Application in Management and Economics. Homewood, Illinois: Richard D. Irwin, 1960.

Bradley, Jack J. and James McClelland. Basic Statistical Concepts, A Self-Instructional Text. Chicago: Scott, Foresman and Company, 1963.

Bruning, James L. and B. L. Kintz. Computational Handbook of Statistics. Glenview, Illinois: Scott, Foresman and Company, 1968.

Bryant, Edward Clark. Statistical Analysis. New York: McGraw-Hill Book Company, 1960.

Buros, Oscar Krisin. Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.

Chase, Clinton I. Elementary Statistical Procedures. New York: McGraw-Hill Book Company, 1967.

Clelland, Richard C. and others. Basic Statistics with Business Applications. New York: John Wiley and Sons, 1966.

Cohen, Lillian. Statistical Methods for Social Scientist, an Introduction. New York: Prentice-Hall, Inc., 1954.

Cornell, Francis Griffith. The Essentials of Educational Statistics. New York: John Wiley and Sons, 1956.

Cowden, Dudley Johnstone. Statistical Methods in Quality Control. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

Cox, David Roxbee. Planning of Experiments. New York: John Wiley and Sons, 1958.

Crouch, Ralph. Finite Mathematics with Statistics for Business. New York: McGraw-Hill Book Company, 1968.

Diamond, Solomon. Information and Error; an Introduction to Statistical Analysis. New York: Basic Books, 1959.

Dixon, Wilfrid Joseph and Frank J. Massey. Introduction to Statistical Analysis. New York: McGraw-Hill Book Company, 1957.

Dornbusch, Sanford M. A Primer of Social Statistics. New York: McGraw-Hill Book Company, 1955.

Downie, Norville Morgan and R. W. Heath. Basic Statistical Methods. New York: Harper and Row, 1959.

Dubors, Edward N. Essential Methods in Business Statistics. New York: McGraw-Hill Book Company, 1964.

Edwards, Allen Louis. Statistical Methods for the Behavioral Sciences. New York: Holt, Rinehart, and Winston, 1954.

Ekeblad, Frederick A. The Statistical Method in Business; Applications of Probability and Influence to Business and other Problems. New York: John Wiley and Sons, 1962.

Elzey, Freeman F. A Programmed Introduction to Statistics. Belmont, California: Wadsworth Publishing Company, 1966.

Ernst, Morris Leopold. The Comparative International Almanac. New York: The Macmillan Company, 1967.

Ferguson, George Andrew. Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Company, 1966.

Fisher, Ronald Aylmer. The Design of Experiments. New York: Hafner, 1951.

________. Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd, 1950.

Franzblau, Abraham Norman. A Primer of Statistics for Non-Statisticians. New York: Harcourt, Brace, and World, 1958.

Freeman, Linton C. Elementary Applied Statistics for Students in Behavioral Science. New York: John Wiley and Sons, 1965.

Freund, John E. and Frank J. Williams. Dictionary/Outline of Basic Statistics. New York: McGraw-Hill Book Company, 1966.

Games, Paul Allen and George Klare. Elementary Statistics; Data Analysis for the Behavioral Sciences. New York: McGraw-Hill Book Company, 1967.

Garlington, Warren King and Helen Shemota. Statistically Speaking. Springfield, Illinois: Thomas, 1964.

Garrett, Henry Edward. Statistics in Psychology and Education. New York: David McKay Company, 1958.

Goldberger, Arthur Stanley. Econometric Theory. New York: John Wiley and Sons, 1964.

Guest, Lester Philip. Beginning Statistics. New York: Crowell Publishing Company, 1957.

Guilford, Joy Paul. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Company, 1956.

Hadley, George. Elementary Statistics. San Francisco: Holden-Day, 1969.

Hagood, Margaret Jarman and Donnie Price. Statistics for Sociologists. New York: Holt, Rinehart and Winston, 1952.

Hammond, Kenneth R. Introduction to the Statistical Method; Foundations and Use in the Behavioral Sciences. New York: Knopf, 1962.

Hill, Joseph E. and August Kerber. Models, Methods and Analytical Procedures in Education Research. Detroit: Wayne State University Press, 1967.

Hilsenrath, Joseph and others. OMNITAB: A Computer Program for Statistical and Numerical Analysis. Washington, D. C.: U. S. Government Printing Office, 1966.

Holl, Paul Gerhard. Elementary Statistics. New York: John Wiley and Sons, 1966.

Hope, K. Elementary Statistics; A Workbook. Oxford, New York: Pergamon Press, 1967.

Huff, Darrell. How to Lie with Statistics. New York: Norton, 1954.

Kendall, Maurice G. and Alison S. Doig. Bibliography of Statistical Literature. Edinburgh: Oliver and Boyd, 1962.

Kendall, Maurice G. A Dictionary of Statistical Terms. New York: Hafner, 1967.

Kurtz, Albert Kenneth and Harold A. Edgerton. Statistical Dictionary of Terms and Symbols. New York: Hafner, 1967.

Loveday, Robert. A First Course in Statistics. Cambridge, England: University Press, 1958.

McCarthy, Philip John. Introduction to Statistical Reasoning. New York: McGraw-Hill Book Company, 1957.

McCollough, Celeste and Loch Atta. Introduction to Descriptive Statistics and Correlation; a Program for Self-Instruction. New York: McGraw-Hill Book Company, 1965.

Mandel, John. The Statistical Analysis of Experimental Data. New York: Interscience Publishers, 1964.

Mark, Mary Louise. Statistics in the Making; a Primer in Statistical Survey Method. Columbus, Ohio: Bureau of Business Research, College of Commerce and Administration, Ohio State University, 1958.

Mood, Alexander McFarlane and Franklin A. Graybill. Introduction to the Theory of Statistics. New York: McGraw-Hill Book Company, 1963.

Masteller, Frederick Rourke and George B. Thomas, Jr. Probability with Statistical Application. Reading, Massachusetts: Addison-Wesley Publishing Company, 1961.

Ostle, Bernard. Statistics in Research; Basic Concepts and Techniques for Research Workers. Ames, Iowa: State College Press, 1954.

Reichmann, William John. Use and Abuse of Statistics. New York: Oxford University Press, 1962.

Schmid, Calvin Fisher. Handbook of Graphic Presentation. New York: The Ronald Press Company, 1954.

Schoer, Lowell A. An Introduction to Statistics and Measurement: A Programmed Book. Boston: Allyn and Bacon, 1966.

Senders, Virginia L. Measurement and Statistics; a Basic Text Emphasizing Behavioral Science Applications. New York: Oxford University Press, 1958.

Spear, Mary Eleanor. Charting Statistics. New York: McGraw-Hill Book Company, 1952.

Statistics Services. A Subject Guide to Data on Industrial Business, Social, Educational, Financial and other Topics for the U. S. and Selected Foreign Countries. Detroit: Gale Research and Book Tower, 1965.

Sterling, Theodor D. and Symour Pollack. Introduction to Statistical Data Processing. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.

Underwood, Benton J., Janet Spence, and others. Elementary Statistics. New York: Appleton-Century-Crofts, 1968.

Walker, Helen Mary. Mathematics Essentials for Elementary Statistics; a Self-Teaching Manual. New York: Holt, Rinehart and Winston, 1951.

Wallis, Wilson Allen and Harry Roberts. Statistics, a New Approach. Glencoe, Illinois: Free Press, 1956.

Weinberg, George H. and John A. Schumaker. Statistics; an Intuitive Approach. Belmont, California: Brooks-Cole Publishing Company, 1962.

Wert, James Edwin, Charles Nardt, and J. Stanley Ahmann. Statistical Methods in Educational and Psychological Research. New York: Appleton-Century-Crofts, 1954.

Williams, Frederick. Reasoning with Statistics; Simplified Examples in Communications Research. New York: Holt, Rinehart and Winston, 1968.

Young, Robert K. and Donald J. Veldman. Introductory Statistics for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1965.

Yuker, Harold E. A Guide to Statistical Calculations. New York: Putman, 1958.

Yule, George Udny and M. S. Kendall. An Introduction to the Theory of Statistics. London: C. Griffin, 1950.

Yumane, Tars. Statistics; An Introductory Analysis. New York: Harper and Row, 1967.

Zeiscl, Hans. Say It With Figures. New York: Harper and Row, 1968.

APPENDIX C

SUGGESTED BIBLIOGRAPHY

COMPUTER APPLICATION

Andree, Richard V. Computer Programming and Related Mathematics. New York: John Wiley and Sons, 1967.

Arnold, Robert R., Harold C. Hill, and Aylmer Nichols. Introduction to Data Processing. New York: John Wiley and Sons, 1966.

________. Modern Data Processing. New York: John Wiley and Sons, 1969.

Beshers, James M. (ed.). Computer Methods in the Analysis of Large Scale Social Systems. Cambridge, Massachusetts: Massachusetts Institute of Technology Press, 1968.

Borko, Harold (ed.). Computer Applications in the Behavioral Sciences. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1962.

Brayton, Jack. Computer Self-Taught through Experiments. Indianapolis: H. W. Sams, 1966.

Chorafas, Dimitris N. Programming Systems for Electronic Computers. London: Butterworths, 1962.

Colman, Harry L. and Clarence Smallwood. Computer Language; an Autoinstructional Introduction to Fortran. New York: McGraw-Hill Book Company, 1962.

________. Computer Yearbook and Directory. Detroit: American Data Processing, 1966.

________. Computers and Education; a Workshop Conference at University of California, Irvine. New York: McGraw-Hill Book Company, 1967.

Cushman, Thomas J. and William J. Keip. Data Processing, a Text and a Project Manual. New York: McGraw-Hill Book Company, 1967.

Davis, Gordon Bilter. Computer Data Processing. New York: McGraw-Hill Book Company, 1969.

Fisher, F. Peter and George F. Swindle. Computer Programming Systems. New York: Holt, Rinehart and Winston, 1964.

Flores, Ivan. Computer Programming. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966.

________. Computer Software; Programming Systems for Digital Computers. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1965.

Green, Bert F. Digital Computers in Research, an Introduction for Behavioral and Social Scientists. New York: McGraw-Hill Book Company, 1963.

Greenwood, Frank and Edwin M. Danziger. Computer System Analysts: Problems of Education, Selection and Training. New York: American Management Association, Administrative Services Division, 1967.

Haag, James N. Comprehensive Standard Fortran Programming. New York: Hayden Book Company, 1969.

Haga, Enoch J. Understanding Automation; A Data Processing Curriculum Guide and Reference Text. Elmhurst, Illinois: Business Press, 1965.

Hassitt, Anthony. Computer Programming and Computer Systems. New York: Academic Press, 1967.

Huskey, Harry D. and Granino A. Korn. Computer Handbook. New York: McGraw-Hill Book Company, 1962.

International Federation for Information Processing. Information Processing, 1965: Proceedings of International Federation for Information Processing Congress, 1965. Washington, D. C.: Spartan Books, 1965.

Leeds, Herbert D. and Gerald M. Weinberg. Computer Programming Fundamentals. New York: McGraw-Hill Book Company, 1966.

Lehman, Richard S. and Daniel E. Bailey. Digital Computing Fortran IV and Its Applications in Behavioral Science. New York: John Wiley and Sons, 1968.

Lytel, Allan Herbert. ABC's of Computer Programming. Indianapolis: H. W. Sams, 1962.

Nathan, Robert and Elizabeth Haynes. Computer Programming Handbook; a Guide for Beginners. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1961.

________. Computer Programming Handbook; a Guide for Beginners. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1962.

Sippl, Charles. Computer Dictionary and Handbook. Indianapolis: H. W. Sams, 1966.

Smith, Karl Ulrich and Margaret Foltz Smith. Cybernetic Principles of Learning and Educational Design. New York: Holt, Rinehart and Winston, 1966.

Stein, Marvin L. and William D. Munro. Computer Programming, A Mixed Language Approach. New York: Academic Press, 1964.

Sterling, Theodor D. and S. V. Pollack. Introduction to Statistical Data Processing. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.

Swanson, Robert W. An Introduction to Business Data Processing and Computer Programming. Belmont, California: Dickerson Publishing Company, 1967.

Technical Education and Management, Inc., New York. Computer Basics. Indianapolis: H. W. Sams, 1962.

PERIODICALS

Advances in Computers

Business Automation

Computer Decisions

Computer Digest

Computer Journal

Computing Reviews

Data Processing Digest

Data Processing for Education

Data Processor

EDP Analyzer

IBM Journal of Research and Development

IBM Systems Journal

Journal of Data Management

Journal of Educational Data Processing

Management Services