INVESTIGATION OF PREFERENCE DIFFERENCES AND THEIR EFFECTS ON THE SECONDARY EDUCATION SYSTEM OF ATLANTA IN 1985

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

By

Wesley Leroy Hamm

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Operations Research

Georgia Institute of Technology

June, 1974

INVESTIGATION OF PREFERENCE DIFFERENCES AND THEIR EFFECTS ON THE SECONDARY EDUCATION SYSTEM OF ATLANTA IN 1985

Approved:

Stuart J. Deutsch, Chairman

Cecil G. Johnson

Ray L. Sweigert, Jr., Director of Atlanta Assessment Project

H. M. Wadsworth

Date approved by Chairman: 5/24/74

ACKNOWLEDGMENTS

I express my sincere appreciation and indebtedness to the erudite Professor Stuart J. Deutsch. His expert guidance and timely encouragement greatly enhanced this research effort.

The contributions of Professor Cecil G. Johnson, Dr. Ray L. Sweigert, Jr., and Dr. H. M. Wadsworth in bringing this research to a successful completion are gratefully acknowledged.

Thanks are extended to the personnel of the Atlanta Assessment Project for providing the data base for this research.

Thanks are also due to Mr. Bruce Schmeiser for his assistance on a portion of the computer programming required for the data analysis.

I would be remiss if I failed to thank my mother, Velma G. Hamm, for instilling in me the desire to reach this rung of the educational ladder.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

LIST OF TABLES v

LIST OF ILLUSTRATIONS vi

SUMMARY vii

Chapter

I. INTRODUCTION 1

1.0 Statement of the Problem 1
1.1 Brief History Leading to the Problem 1
1.2 Purpose of the Research 3
1.3 Brief Overview of the Research 4

II. LITERATURE REVIEW 6

2.0 Introduction 6
2.1 Goals of Secondary Education 7
2.1.1 Goal Development and Evaluation . . . 7
2.1.2 Curriculum Development and Evaluation 11
2.1.3 Interest-Oriented Courses 13
2.1.4 Career Education 15
2.2 Statistical Analysis of Ordinal Data 18
2.2.1 Correlational Methods 18
2.2.2 Statistical Tests for Two Samples . . 26
2.2.3 Statistical Tests for K Samples . . . 29
2.2.4 Summary 31

III. EXPLANATION OF DATA BASE 33

3.0 Introduction 33
3.1 Delphi Studies 34
3.2 Goals for Education in Atlanta, 1985 54

IV. OBJECTIVES 56

4.0 Introduction 56
4.1 Category One Hypotheses 58
4.2 Category Two Hypotheses 60

Chapter		Page
V.	METHODOLOGY.	62
	5.0 Introduction.	62
	5.1 Ford's Procedure.	63
	5.1.1 Background	63
	5.1.2 Advantages	66
	5.1.3 Application to Data Base	67
	5.2 Randomization Distribution.	69
	5.3 Comparative Example	70
VI.	DELPHI TECHNIQUE--CATEGORY ONE HYPOTHESIS TESTS.	79
	6.0 Introduction.	79
	6.1 Application of the Delphi Technique in Studies of Educational Goals.	80
	6.2 Procedure	83
	6.2.1 General Procedure.	83
	6.2.2 Procedure Utilized in Collection of Data Base.	83
	6.3 Determination of Convergence.	85
	6.3.1 Methods Normally Employed.	85
	6.3.2 Recommended Method	86
	6.4 Comparative Example	93
	6.5 Results of Category One Hypothesis Tests. .	101
	6.6 Summary	103
VII.	CATEGORY TWO HYPOTHESIS TESTS.	105
	7.0 Introduction.	105
	7.1 Comparative Example	106
	7.2 Results of Category Two Hypothesis Tests. .	111
VIII.	CONCLUSIONS.	125
APPENDIX A.	. . . . .	130
APPENDIX B.	. . . . .	132
APPENDIX C.	. . . . .	133
APPENDIX D.	. . . . .	134
BIBLIOGRAPHY.	. . . . .	143

LIST OF TABLES

Table		Page
1.	Taxonomic Goal Areas.	36
2.	Goals Recommended by Community Leaders.	38
3.	Goals Recommended by Students	41
4.	Goals Recommended by Educators.	42
5.	Goals for Education in Atlanta, 1985, as Seen by Community Leaders, Educators and Students. . .	43
6.	Goal Area 2.2: Student Delphi Study.	94
7.	Results of Category One Hypothesis Tests.	102
8.	Forming Relationships with Others: Community Leaders Versus Students	108
9.	Goal Areas in Which the Category Two Null Hypothesis Was Rejected	112
10.	Results of Category Two Hypothesis Tests.	117

LIST OF ILLUSTRATIONS

Figure Page

1. Example Problem: Randomization Versus t Distribution. 73

2. Consistency in Preference Graph 89

3. Career Development: Student Delphi Study, Rounds One and Two. 95

4. Career Development: Student Delphi Study, Rounds Two and Three. 97

5. Forming Relationships with Others: Community Leaders Versus Students 109

SUMMARY

The purpose of this research is to determine the value of the information in each goal area obtained from the Delphi studies and, within those goal areas containing information of value, to investigate the differences in preference between community leaders, students and educators. Additionally, the specific areas in which preference differences exist will be identified and their probable effects will be noted. By examining the relative importance attached to specific goals, insight will be provided into the problem of what type of end product the respondents believe the school should produce.

The methodology developed allows the researcher to make an exact probability assessment concerning the degree of consistency in preference achieved within a category of judges during a Delphi study, i.e., the value of the information within a category of judges obtained from a Delphi study. If a sufficient amount of consistency between successive questionnaires within a category of judges is obtained, the information contained in the last questionnaire for that category of judges may be used in future statistical tests. Ford's procedure is employed to achieve a weight for each goal rated on the questionnaires. Although the transformation of the Ford weights to ranks is useful in

itself, much information concerning the relative magnitudes of the weights is lost by their conversion. By allowing the goals to remain associated with the actual weights generated by Ford's procedure, a clearer indication of the degree of importance associated with each goal for a particular questionnaire of a category of judges is obtained. The sample distribution utilized in the hypothesis tests involving Ford weights is the randomization distribution which was developed from Fisher's Principle of Randomization.

The categories of hypothesis tests used in the research are: (a) category one, which determines whether a category of judges' preferences changed over time, i.e., whether the preferences of the community leaders or the students or the educators changed between rounds one and two or between rounds two and three; and (b) category two, which determines whether the preferences between categories of judges differ significantly, i.e., whether community leaders and students, community leaders and educators or students and educators differ significantly in their preferences.

The category one hypothesis tests indicated that the Delphi technique was successful in producing consistent preferences in eight out of 14 goal areas in each of the three Delphi studies conducted. These results indicate that further investigation is required in the six goal areas that did not contain consistent preferences before any

category two hypothesis tests are conducted involving them.

The category two hypothesis tests, conducted only in those goal areas in which the Delphi technique was successful in obtaining consistent preferences, indicated that community leaders and students are closest in agreement with respect to their preferences for educational goals. The community leaders and educators differed significantly in two goal areas; and educators and students differed significantly in four goal areas.

CHAPTER I

INTRODUCTION

1.0 Statement of the Problem

The development of secondary level educational goals for the Atlanta of 1985 is one of the tasks of the Atlanta Assessment Project. To accomplish this task, three separate Delphi studies were conducted during the latter part of 1972. This research will utilize the data collected by the Atlanta Assessment Project to assess the value of the information derived from these Delphi studies. Methods for accomplishing this assessment vary from ad hoc procedures to detailed statistical analysis. The validity of the procedures usually employed is dependent on the criterion of measurement adopted. A second problem involved is the analysis of differences in the perceived importance of educational goals. Specific areas in which differences in preference exist will be identified and the effects of these differences will be explored.

1.1 Brief History Leading to the Problem

Traditionally, the goals of our secondary educational process have been developed by scholars in the field of education. The Atlanta Assessment Project, which began in March 1972, has taken a unique departure from this tradition.

The overall objective of the project is: "To determine the progress of twelfth graders, 17-year-old students at lower grades, and 17-year-old nonstudents, in the area served by the Atlanta Public Schools, in respect to mastery of the Goals for Education in Atlanta at a minimal level necessary for successfully coping with life in the Atlanta of 1985." In this project professional, technical, managerial and community leaders as well as high school leaders are given equal weight with high school teachers, counselors, principals and other administrators in the development and ranking of secondary educational goals. The Delphi technique was selected because it was believed that each of the three groups represented a distinct area of expertise with respect to goal development and evaluation in secondary education. The project began with a list of 86 goals that were derived from the Goals for Education in Georgia which had been developed by the State Advisory Commission on Educational Goals and approved by the State Board of Education in 1970. The intent of the Delphi studies was to determine the applicability of these goals to the Atlanta area in the aforementioned time period; to determine what, if any, additional goals should be included; and to determine the relative importance of each of the goals [3].

The respondents were asked to judge the degree of future importance of the goals on an integer scale from one to six. The verbal descriptions associated with the ratings

were: 6-of the highest importance; 5-of very high importance; 4-of high importance; 3-of medium importance; 2-of some importance; and 1-of little or no importance. An important part of the surveys was to invite each respondent to suggest possible additional goals that he or she felt were necessary and should be included. Each suggested goal was reviewed in terms of three criteria: (a) Did the suggestion communicate, i.e., was it understandable?; (b) Did the suggestion concern student learning outcomes as opposed to institutional goals?; and (c) Did the suggestion contain a substantive element appreciably different from those in the initial set of goals? Thirty-five goals were added to the original 86 goals.

This research will utilize the preferences of individuals obtained by the Atlanta Assessment Project. These preferences are ordinal in nature and are contained on 10,000 data cards.

1.2 Purpose of the Research

The purpose of this research is to determine the value of the information in each goal area obtained from the Delphi studies and, within those goal areas containing information of value, to investigate the preference differences between community leaders, students and educators. Additionally, the specific areas in which inconsistencies exist will be identified and their probable effects will be

noted. By examining the relative importance attached to specific goals, insight will be provided into the problem of what type of end product the respondents believe the school should produce.

1.3 Brief Overview of the Research

Chapter II highlights the development and evaluation of career education, interest-oriented courses, curriculum and educational goals. The contrasting views that currently exist in many of these areas are examined. Also in Chapter II, the historical development of nonparametric correlational methods and the pertinent nonparametric statistical tests used in the analysis of ordinal data are presented.

A description of the data base provided for this research is presented in Chapter III. The taxonomic goal areas and the goals for education in Atlanta, 1985, are also delineated in this chapter.

The objectives of the research and the two categories of hypothesis tests that will be utilized in meeting the objectives are defined in Chapter IV. The implications to the education process with the overall objectives are also discussed in this chapter.

The methodology proposed in this research and the background, advantages and applicability of Ford's procedure are presented in Chapter V. Fisher's Principle of Randomization and the method of analysis derived from it are also

discussed in this chapter. A comparative example is presented to illustrate the proposed methodology and to compare it with other methods of analysis commonly used.

The definition, objectives and applications of the Delphi technique are presented in Chapter VI. The general procedure and the procedure used by the Atlanta Assessment Project are outlined in this chapter. The methods and rationale used to determine whether the Delphi technique has been successful in producing consistent preferences within each category of judges are also presented, i.e., the method used in determining the value of the information derived from the Delphi technique. The chapter concludes with a comparative example and the results obtained from applying the preference consistency criteria to the data base used in this research.

The determination of specific goal areas in which differences in preference between categories of judges exist is discussed in Chapter VII. A comparative example and the results obtained during this research are also presented in this chapter.

The final chapter, Chapter VIII, states the conclusions arrived at with respect to the proposed methodology, the Delphi technique and the goal areas in which significant differences in preference between categories of judges were observed. The implications of these preference differences are also discussed.

CHAPTER II

LITERATURE REVIEW

2.0 Introduction

In a research problem of this nature it quickly becomes apparent that there are two distinct areas involved. The first area is that dealing with the goals of secondary education, and the second area is that of statistical analysis of individual and group preferences. In the area of goals of secondary education, the topics of goal development and evaluation, curriculum development and evaluation, interest-oriented courses and career education will be discussed. It is immediately recognizable that there are contrasting views and opinions on many of these topics; with respect to the data collected and the Atlanta of 1985, a number of these issues will be examined in subsequent chapters. Normally, the statistical analysis of individual and group preferences would include both parametric and nonparametric methods. However, since the data for this research are based on an ordinal scale, parametric methods are not appropriate. Those nonparametric statistical methods applicable to ordinal data will be examined in the statistical analysis portion of this chapter.

2.1 Goals of Secondary Education

2.1.1 Goal Development and Evaluation

The development and evaluation of the goals and objectives of secondary education have received much attention over the past forty years. Despite the attention received, it appears that there is as much, if not more, debate and disagreement today as there was at any time in the past. As Palmer [40] recently noted, educators are not in agreement as to the need for educational objectives and many of those who do acknowledge the need do not use the objectives in their classes. He suggested that until those who write educational objectives are forced into the harsh world of reality and compromise, they will continue to sit in their ivory towers insisting on the unattainable from those they are trying to convince.

In 1948, at the American Psychological Association Convention in Boston, the seeds were planted for what was to become the Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. The college examiners attending this convention felt that the purposes of the taxonomy should be: (a) to help teachers, administrators, professional specialists and research workers who deal with curricular and evaluation problems to discuss their problems with greater precision; (b) to facilitate the exchange of information about curricular developments and evaluation devices; (c) to suggest the kinds of objectives that can be included in a curriculum;

(d) to help teachers and others gain a perspective on the emphasis given to certain behaviors by a particular set of educational plans; and (e) to help curriculum builders to specify objectives so that it becomes easier to plan learning experiences and prepare evaluation devices. In constructing the taxonomy, every effort was made to avoid value judgments about objectives and behaviors. The boundaries between categories were chosen so as to coincide as closely as possible with the distinctions teachers make in planning curricula or in choosing learning situations.

In 1956 the Taxonomy of Educational Objectives, Handbook I: Cognitive Domain was completed. The cognitive domain was written first because most of the work in curriculum development dealt with this domain, and the clearest definitions of objectives were phrased as descriptions of student behavior in this domain [5]. The cognitive domain deals with those objectives which emphasize remembering or reproducing something which has presumably been learned, as well as objectives which involve the solving of some intellective task for which the individual has to determine the essential problem then rearrange given material or combine it with ideas, methods or procedures previously learned. The taxonomy delineates six major classes of objectives, each class making use of and being built on the behavior found in the preceding classes. Since the classes were arranged from simple to complex, a hierarchical arrangement

resulted. A condensed version of this portion of the taxonomy is located in Appendix A.

In 1964 Krathwohl, Bloom and Masia [31] published their taxonomy on the affective domain. The affective domain deals with those objectives which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection. Affective objectives vary from simple attention to selected phenomena to complex with internally consistent qualities of character and conscience. The taxonomy was developed using the same structure scheme that was used in the taxonomy for the cognitive domain. A condensed version of the affective domain of the taxonomy is located in Appendix B.

Psychomotor domain objectives are those which emphasize some muscular or motor skill, some manipulation of material and objects, or some act which requires a neuromuscular coordination. The authors of the taxonomies for the cognitive and affective domains have stated, "Although we recognize the existence of the psychomotor domain, we find so little done about it in secondary schools or colleges, that we do not believe the development of a classification of these objectives would be very useful at present." However, in 1966, Simpson [45] stated, "We believe that the psychomotor domain has relevance for education in general as well as for such areas of specialization as industrial education, agriculture, home economics, business education, music, art and physical education." Educators in her own

field of vocational and technical education have a keen interest in the development of a classification system for educational objectives in this domain because many technical jobs require a high degree of ability and skill in the psychomotor domain as well as in the cognitive and affective areas. In this same year Simpson [45] developed a classification system for educational objectives, psychomotor domain. She admitted that the form was tentative, flexible and incomplete. Simpson and other vocational education leaders agreed that development of a taxonomy for the psychomotor domain was basic to other research on the development of psychomotor abilities and skills needed in vocational technical education. The proposed taxonomy for the psychomotor domain is located in Appendix C.

Although the development of the taxonomies has greatly enhanced the communicability and hierarchical nature of educational objectives, they have not succeeded in uniting educators. There are basically two schools of thought concerning the objectives of education. One school of thought believes that the taxonomic structure presented above is a vital and necessary prerequisite for curriculum development and evaluation. Palmer and others believe that a taxonomic structure is essential; but, until those who write educational objectives and those who use educational objectives are willing to compromise, the taxonomies will remain a theoretical model instead of a working model. A

second school of thought is that educators should be more concerned with general purposes and goals than with detailed outcomes [12]. That is, the taxonomies deal with abstract intellectualism in a well organized, definitive manner; however, the broader concept of educating the whole person is neglected. Since the Atlanta Assessment Project used the taxonomies mentioned above in classifying the educational goals that will be utilized in this research, no attempt to restructure or reclassify the goals will be undertaken in this research.

2.1.2 Curriculum Development and Evaluation

The subject of curriculum development has been of major concern for many years. As noted previously, educators do not agree on the validity or the use of the objectives stated in the taxonomies. Herein lies a stumbling block for curriculum development. Also, as Potter [42] noted, "It would seem that some of the chief assets of the country as a whole are among the chief factors that hinder its advance in education--its wealth, its technological progress, its love of freedom." There are also differing views on the function of curriculum. Hogan [25] believed that the function of curriculum is the same as that of the wall in handball. Without it, the game that takes place mostly on the open court would not be possible. Additionally, although everyone agrees that our high schools are in difficulty, arguments immediately arise over the source and the size of

the trouble, as well as what is wrong [37].

The taxonomies have provided for communicability among educators and researchers, and they have proven to be a valuable tool in certain areas of curriculum development. Finch [17] noted, "The taxonomies of the cognitive, affective and psychomotor domains provide a systematic, comprehensive approach to curriculum development in business education and are applicable in revising, reviewing and establishing objectives for lessons, courses and programs in business education."

There are those who disagree with the concept of detailed educational objectives. Ebel [12] proposed that it is more urgent for educators to reach agreement on their general purposes and goals than to specify in detail the outcomes they seek. He also noted that insistence on detailed statements of educational objectives is questionable. He defended the proposition that teachers should be more concerned with developing a pupil's cognitive resources than with changing his behavior.

In 1969 Goodlad [23] noted that, after a flurry of activity during the late thirties and early forties, curriculum evaluation laid virtually dormant during the fifties and early sixties. He also noted that little has been added to the knowledge of deriving educational objectives since the first taxonomy was published in 1956. His opinion was that curricular theory with exploratory and predictive power

was nonexistent.

To make valid and meaningful curriculum changes, one must first understand those for whom the curriculum is intended and the environment in which it will be presented. In 1970, Wilhelms [55] noted that tomorrow's students will come to school better equipped physically, emotionally and intellectually. As Unruh [51] pointed out recently, "Perhaps the greatest lesson learned from the curriculum reform efforts of the sixties is the need for increased capacity for renewal and responsiveness in curriculum development." Until 1960, there was a high degree of uniformity in the content of courses taught throughout the United States [24]. This was due to the limited number of texts available for a particular course and the same courses were taught at given levels.

2.1.3 Interest-Oriented Courses

Tanner [49] observed that perhaps for the first time the demand for curricular reconstruction is coming from the young--not from the philosophers, the scientists or the politicians. As Foshay [20] noted, "The students want to see themselves as participants in the world they live in, not as apprentices for it. They want the world to be in the school and the school in the world." Quignard [43] stated that one of the problems of secondary education is that the educational system is not answering the cultural and psychological needs of the great mass of young people.

In 1971 Briggs [8] noted that the basic duties of a

secondary school are to teach students to do better the desirable things that they will do anyway and to reveal higher activities and to make them both desired and, to an extent, possible. Walen [53] observed that the school has an increasing need to stretch beyond the walls and to offer real world learning experiences in addition to the academic curriculum.

Many schools have initiated programs that enable the students to take special interest courses or participate in the real world while continuing their academic course work. Fenner [16] described a high school curriculum with forty real world oriented electives that attempt to inject relevance into learning. Ohme [39] noted that the use of interest-centered nine-week electives had a notable change in attitude on the part of both teachers and students. He observed a shift in emphasis from grades to a desire to learn and a shift from apathy to involvement.

At Portland, Oregon's Lincoln High School a program of off-campus enrichment courses for credit has had success [14]. The program invites juniors, sometimes freshmen and sophomores, to move into part-time jobs, volunteer community service, or courses at nearby postsecondary institutions and vocational centers. Besides educational advantages, the program has motivated youngsters, provided more freedom for students and increased their sense of responsibility.

Pepe [41] reported that for the past ten years the

district school of Southbury-Middlebury has offered short-term or mini-courses on Saturday mornings at no charge to the students. The program has met with genuine success, and it involved minimal cost.

Bloom [6] recently reported on Worcester Academy's intern program for juniors and seniors. The internships run the gamut of student interests. Some of the more popular ones are urban affairs, politics, law enforcement and social service. Each intern must read at least one paperback book per week and submit a written evaluation of the book. Bloom noted, "Educators must be keenly aware of students' demands to help in the formulation of their own education, and we must come to respect their belief that education can take place outside of the classroom."

2.1.4 Career Education

The question of career education versus general education has received much attention in recent years. In 1966 Wright [56] advocated not segregating those vocational students from the social life offered by the comprehensive high school. He noted that ninety-five percent of job failures can be attributed to inadequacies in personality development. In 1969 Evans [15] stated, "American education is designed according to one basic objective: to prepare the student for subsequent schooling." He also noted that, in general, curriculum updating has been made with the intent of better preparing students for college. Furthermore, less

than twenty percent of the high school students in the United States are enrolled in vocational education courses. At the time that Mr. Evans's article was written, there were no data at the national level on the number of students completing high school vocational education programs nor on the number of high schools that offer vocational education programs.

In 1972 Marland [36] asserted that every person in school belongs in the "career education" group at some point, whether engaged in preparing to be a surgeon, a bricklayer or a secretary. He observed that traditionally, vocational education has been treated like education's poor cousin and has been typified by dull classes in dull buildings. He also noted that only two out of ten high school students complete college level work. Thus eight out of ten high school students should be receiving some sort of occupational training but only two out of these eight are receiving such training. Hence, as Marland stated, "Fully one-half of our high school students are getting irrelevant, educational pap!"

Thomas Glennan, Director of the newly formed National Institute of Education (NIE), has recognized the need to narrow and classify the role of career education [37]. He presents the following definition of career education: "Career education is the development of knowledge and of special and general abilities to help individuals interact with the economic sector. Learning in this context would

occur in both formal and informal situations which motivate the learner by causing him to experience work directly."

In 1969 Livingston [35] conducted a survey of high school graduates from the Acalanes High School in California to determine, among other things, the degree of satisfaction with their high school education. The following findings were reported: (a) fifty-seven percent of the respondents thought that the school had been of little or no help in their learning how to conduct their own business affairs; (b) fifty-three percent of the respondents thought that the school had been of little or no help in providing information about various occupations; (c) fifty-six percent of the respondents thought that the school had been of little or no help in providing vocationally-oriented courses; and (d) ninety-five percent of the respondents indicated that the school was helpful in providing a good academic background for further education.

In a similar survey, Betz, Engle and Mallinson [4] interviewed a randomly selected group of noncollege-bound high school graduates and obtained some interesting perceptions. Their analysis resulted in four major conclusions: (a) employment-bound, noncollege-oriented students perceive the school, the counselors and other personnel within the school as favoring the college-bound student; (b) counselors were not perceived as being helpful in assisting employment-bound youth to satisfactory vocational decisions; (c) subjects

were unable to articulate meaningful concepts of self; and (d) generally, they did not perceive parents as being at all helpful in resolving personal, educational and vocational problems. Their study substantiated Friedenberg's proposition that a perceptual dichotomy does exist among high school students. Many students tended to see the educational and vocational world divided into two major subgroups: the college-bound (the "Frats") and the noncollege-bound (the "Greasers").

In 1971 Nation's Schools magazine's monthly opinion poll of school superintendents indicated that forty-nine percent thought that the noncollege-bound student was most unsatisfactorily served by their current education program [13]. Ninety percent of the superintendents felt that the college-bound student was best served.

2.2 Statistical Analysis of Ordinal Data

2.2.1 Correlational Methods

The area of statistical analysis dealing with correlational methods had its beginning in 1846. It was in that year that Bravais discovered the mathematical relationships for calculating a correlation coefficient, r, that was later to be named the "product moment" correlation coefficient by Pearson [46]. However, the first person to recognize the possibility of the immense advances that were later made in the measurement of association between two things was

Galton [46]. In 1886, Galton wrote: "The length of the arm is said to be correlated with that of the leg, because a person with a long arm has usually a long leg and conversely."

Rank correlation seems to have had its origin in the method of representing the distribution of a variate by grades or percentiles introduced by Galton in 1888 [22]. In 1906 Spearman [47] experimented with converting absolute measurements into relative ranks and then computing a correlation coefficient, a method which he called "a foot-rule for measuring correlation." He calculated the gain in ranks of a set of data over a set of previously gathered data. He called the sum of gains on the average due to mere chance M and his correlation coefficient R. He calculated the correlation coefficient by the formula

$$R = 1 - \frac{\Sigma(\text{gains})}{M} . \qquad (1)$$

He also noted that in a sample of size n there was a probable error associated with R and this probable error was equal to $.43/(n)^{1/2}$. He related his formula to Pearson's and noted that if the two sets of data are ranks, then Pearson's r is equivalent to

$$r = 1 - \frac{\sum_{i=1}^{n} d_i^2}{m} , \qquad (2)$$

where $m=n(n^2-1)/6$, n = number of observations in each sample and d_i = rank difference between the i^{th} observations. In most texts this rank correlation coefficent is denoted as Spearman's rho (r_s). Spearman's rho is most often used to test the null hypothesis that the two variables under study are not associated in the population and that the observed value of r_s differs from zero only by chance.

The efficiency of the Spearman rank correlation coefficient when compared to the most powerful parametric correlation coefficient, the Pearson r, is about 91 percent.

Kendall [28] introduced a new rank correlation coefficient called τ (tau) in 1938. To calculate τ, the maximum possible score and the actual score must be computed. The maximum possible score in a ranking of n objects is shown to be n(n-1)/2. He gives two rules for computing the actual score, S. The rank correlation coefficient is then

$$\tau = \frac{S}{\frac{1}{2}\, n(n-1)}\ , \quad -1 \leq \tau \leq 1. \tag{3}$$

The sampling distribution of τ under the null hypothesis is known, and therefore is subject to tests of significance. The null hypothesis is that τ differs from zero only by chance, i.e., the two variables are not associated.

Kendall's tau (τ) and Spearman's rho (r_s) have the same power to detect the existence of association in the

population. That is, the sampling distribution of τ and r_s are such that with a given set of data both will reject the null hypothesis (that the variables are unrelated in the population) at the same level of significance. The primary advantage of Kendall's tau lies in its ability to be extended to a partial rank correlation coefficient. A partial rank correlation coefficient is one that measures the degree of association among two variables, x and y, when a third variable, z (on which the association between x and y might logically depend), is held constant. Kendall's tau also has the advantage of having a sampling distribution that is practically indistinguishable from a normal distribution for sample sizes as small as nine.

In 1939, Kendall and Smith [29] first considered the problem of m rankings on n objects. No assumptions were made about the nature of the quality according to which the objects were ranked, other than that ranking is possible. No hypothesis was made about the quality being measurable nor about any underlying frequency distribution. They defined a coefficient of concordance, W, and calculated its value by

$$W = \frac{12S}{m^2(n^3-n)} \quad , \quad 0 \leq W \leq 1. \tag{4}$$

Here, S is the observed sum of squares of the deviations of sums of ranks from the mean value $m(n+1)/2$. They related W

to the average of the m(m-1)/2 Spearman rank correlation coefficients between all pairs of the m rankings. If r'_{av} is this average, then

$$r'_{av} = \frac{\frac{12S}{n^3-n} - m}{m^2-m} \tag{5}$$

or

$$r'_{av} = \frac{mW-1}{m-1} \quad , \quad \frac{-1}{m-1} \leq r'_{av} \leq 1. \tag{6}$$

To test the significance of W, the authors investigated the exact distribution of W (or, more conveniently of S) by permuting the n ranks in all possible ways. They found that for most practical purposes it is most convenient to put

$$z = \frac{1}{2} \log_e \frac{(m-1)W}{1-W} \tag{7}$$

so that z can be tested in Fisher's distribution with $(n-1)-(2/m)=n_1$ and $(m-1)\ (n-1)-(2/m) = n_2$ degrees of freedom. They found that for small values of m and n, a correction factor was needed. If unity is subtracted from S before calculating W and the divisor, $m^2(n^2-n)/12$, is increased by two, the z test will give sufficiently accurate results for n greater than 3.

A high or significant value of W may be interpreted as meaning that the observers or judges are applying essentially the same standard in ranking the N objects under study. Often their pooled ordering may serve as a "standard," especially when there is no relevant external criterion for ordering the objects. A high or significant value of W does not mean that the orderings observed are correct. In fact, they may all be incorrect with respect to some external criterion.

In 1948 Kendall [30] considered testing the null hypothesis that two variables are not associated in some common population and that the observed value of Spearman's rho differs from zero only by chance. He concluded that when the number of observations, N, is greater than 10, the significance of an obtained r_s under the null hypothesis may be tested by

$$t = r_s \sqrt{\frac{N-2}{1-(r_s)^2}} \qquad (8)$$

That is, for N large, the value defined by formula (8) is distributed as Student's t with (N-2) degrees of freedom.

Methods involving rankings were subjected to severe criticism in the early 1900's; however, in 1954 Stuart (48) investigated the correlation between variate-values and ranks and concluded that often there is justification for replacing

the original variates by their ranks. Using some known distributions he showed that there was little loss in efficiency and considerable saving in computation when ranks were used.

In 1956 Siegel [44] authored one of the most comprehensive texts to date on nonparametric methods. He noted that when there are numerous ties present in two rankings, Spearman's rho should be calculated by the formula

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d^2}{2\ \Sigma x^2 \Sigma y^2} \tag{9}$$

where $\Sigma x^2 = \frac{N^3-N}{12} - \Sigma T_x$, $\Sigma_y{}^2 = \frac{N^3-N}{12} - \Sigma T_y$ and ΣT indicates that we sum the various values of T for all the various groups of tied observations in each ranking. He also noted that in the case of numerous ties in two rankings, Kendall's tau should be calculated as

$$\tau = \frac{S}{\sqrt{\frac{1}{2}\ N(N-1)-T_x}\sqrt{\frac{1}{2}\ N(N-1)-T_y}} \tag{10}$$

where $T_x = \frac{1}{2}\ \Sigma t(t-1)$, t being the number of tied observations in each group of ties on the X variable and similarly for T_y. He observed that when the number of observations, N, is greater than 10, Kendall's tau may be considered to be

normally distributed with a mean of zero and a standard deviation of $\sqrt{2(2N+5)/9N(N-1)}$. Thus,

$$z = \frac{\tau - \mu_\tau}{\sigma_\tau} = \frac{\tau}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} \tag{11}$$

is approximately normally distributed with zero mean and unit variance.

In 1964 Lieberson [34] addressed some of the limitations involved in the application of certain nonparametric coefficients of correlation. He noted that most sociological statistics textbooks fail to mention any of the limitations or underlying assumptions associated with Kendall's tau and Spearman's rho. In particular, all monotonic curvilinear functions are suitably measured by tau and rho; however, if the function is nonmonotonic, then their coefficients will underestimate the existing degree of association. He outlined two characteristics that can be used to identify nonmonotonic rank-order functions. The first characteristic is found using a modified "runs" test. Suppose we have two variables, X and Y, and their associated rankings. First, place the X variable in natural order, i.e., 1, 2, 3,..., N. Then, starting with the Y rank opposite X_1, mark each increase or decrease in the following Y rank over the preceding one with + or -, respectively. Let n_1 = the number of plus signs, n_2 = the number of minus signs, and V = the

number of runs of consecutive signs. If the correlation coefficient and the value of V are small, then the function is nonmonotonic. The second characteristic is based on a comparison of runs. We follow the procedure outlined above for both rankings and then compare the two resulting values of V. If and only if a strong nonmonotonic function exists, the number of runs obtained with each ordering will be radically different. This second approach has the advantage of permitting one to determine the applicability of rank-order coefficients prior to their calculation.

2.2.2 Statistical Tests for Two Samples

Wilcoxon's [54] matched-pairs signed-ranks test has proven to be the most useful test for the behavioral scientist. The test is most generally used to test the null hypothesis of no difference in two treatments. A difference in score, d_i, is computed for each matched pair and the resulting d_i's are ranked without regard to sign with the smallest d_i receiving a rank of one. For small samples, N less than 26, let T be the smaller sum of like-signed ranks, then a table of tabulated values is consulted [44]. If an observed T value is less than or equal to the tabulated value, then the null hypothesis is rejected. For large samples, N greater than 25, Siegel [44] noted that T is practically normally distributed with a mean of $N(N+1)/4$ and a variance of $N(N+1)(2N+1)/24$. Thus,

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{N(N+1)}{4}}{\frac{\sqrt{N(N+1)(2N+1)}}{(24)^{1/2}}} \tag{12}$$

is approximately normally distributed with zero mean and unit variance.

The Mann-Whitney U test [44], used to test whether two independent groups have been drawn from the same population, is one of the most powerful of the nonparametric tests. It is a most useful alternative to the parametric t test when the researcher wishes to avoid the t test's assumptions or when the measurement in the research is weaker than interval scaling.

Let N_1 equal the number of observations in the smaller of two independent groups, and let N_2 equal the number of observations in the larger group. To apply the U test, we first combine the observations from both groups and rank these in order of increasing size. In this ranking, algebraic size is considered, i.e., the lowest ranks are assigned to the largest negative numbers, if any. Now focus on one of the groups, say the group with N_1 observations. The value of U (the test statistic) is given by the number of times that a rank in the group with N_2 observations precedes a rank in the group with N_1 observations in the ranking. A table of critical values is consulted to determine if the test statistic U is significant.

For fairly large values of N_1 and N_2, the counting method of determining the value of U may be rather tedious. An alternative method, which gives identical results, is to assign the rank of one to the smallest observation in the combined $(N_1 + N_2)$ group of observations and to assign rank two to the next smallest observation, etc. Then

$$U = N_1N_2 + \frac{N_1(N_1+1)}{2} - R_1 \tag{13}$$

or, equivalently,

$$U = N_1N_2 + \frac{N_2(N_2+1)}{2} - R_2, \tag{14}$$

where R_1 = the sum of the ranks assigned to the group whose sample size is N_1 and R_2 = the sum of the ranks assigned to the group whose sample size is N_2. Formulas (13) and (14) yield different values of U. It is the smaller of these values that is used as the test statistic. A table of critical values is consulted to determine if U is significant.

When $N_2 > 20$, the significance of an observed value of U may be determined by the use of the standard normal statistic z as

$$z = \frac{U-\mu_u}{\sigma_u} = \frac{U - \frac{N_1N_2}{2}}{\sqrt{\frac{(N_1)(N_2)(N_1+N_2+1)}{12}}}. \tag{15}$$

If the Mann-Whitney test is applied to data which might properly be analyzed by the most powerful parametric test, the t test, its power-efficiency approaches 95.5 percent as N increases; and it is close to 95 percent even for moderate-sized samples. It is therefore an excellent alternative to the t test; and, of course, it does not have the restrictive assumptions and requirements associated with the t test.

The randomization test for two samples is a useful and powerful nonparametric technique when N_1 and N_2 are small. The historical development of the test and the assumptions required are discussed in Chapter V.

2.2.3 Statistical Tests for K Samples

In 1940 Friedman [21] investigated various alternative tests of significance for the problem of K rankings. He showed that if we have a table of (KxN) ranks, then as K becomes large,

$$\chi_r^2 = \frac{(N-1)(\text{Column Sum of Squares})}{(N^3-n)/12} \qquad (16)$$

is approximately distributed as χ^2 with (N-1) degrees of freedom under the null hypothesis that the column means are all equal. Relating χ_r^2 to Kendall's coefficient of concordance W, he concluded that

$$\chi_r^2 = K(N-1)W \tag{17}$$

would be tested by the χ^2 distribution tables. In general, for K greater than seven and N less than six, the z test with continuity corrections developed by Kendall and Smith [29], should be employed. The χ^2 test is appropriate for all other values.

In 1952 Kruskal and Wallis [32] developed a statistic for testing the null hypothesis that K samples of ranked data were drawn at random from a common population or form identical populations. In the computation of the Kruskal-Wallis test statistic H, each of the N observations are replaced by ranks. That is, all of the scores from all of the K samples combined are ranked in a single series. The smallest score is replaced by rank 1, the next to smallest by rank 2, etc. When this has been done, the sum of the ranks in each sample (column) is found. The Kruskal-Wallis test determines whether these sums of ranks are so disparate that they are not likely to have come from samples which were all drawn from the same population. The test statistic H may be written as

$$H = \left(\frac{12}{n(n+1)}\right)\left(\sum_{j=1}^{k} \frac{T_j}{n_j}\right) - 3(n+1), \tag{18}$$

where K = number of groups, n_j = number in the j^{th} group, n = total number of observations, and T_j = sum of ranks or rank total for the j^{th} group. The statistic may also be expressed as

$$H = \frac{\text{Column Sum of Squares}}{(n^3-n)/12(n-1)}. \qquad (19)$$

The test assumes that the variable under study has an underlying continuous distribution. It also requires at least an ordinal measurement of that variable. They noted that under the null hypothesis, provided the number of observations in each group is not too small, H is approximately distributed as a Chi-Square with (K-1) degrees of freedom. They observed that the test may be most useful in testing differences in means, without the necessity of assuming homogeneity of variance, since the H test may be relatively insensitive to differences in variances. Siegel [44] notes that the Kruskal-Wallis test seems to be the most efficient of the nonparametric tests for K independent samples. It has a power-efficiency of 95.5 percent, when compared to the F test.

2.2.4 Summary

Although the correlational methods reviewed in section 2.2.1 are applicable to the ordinal data base in this research, they were not utilized because the resolution

provided by the method of analysis employed in this research and described in Chapter V would have been lost. They were presented as a portion of the literature review for completeness only.

The methods which will be examined in this research include the Wilcoxon matched-pairs signed-ranks test, the Mann-Whitney U test, the Kruskal-Wallis test and the randomization test. As a means of comparing the results of these nonparametric tests with those that would have been obtained if parametric tests had been used, the probability associated with the appropriate test statistic from the t distribution will also be given for each of the comparative examples.

CHAPTER III

EXPLANATION OF THE DATA BASE

3.0 Introduction

As noted in Chapter I, the overall objective of the Atlanta Assessment Project is: "To determine the progress of twelfth graders, 17-year-old students at lower grades, and 17-year-old nonstudents, in the area served by the Atlanta Public Schools, in respect to mastery of the Goals for Education in Atlanta at a minimal level necessary for successfully coping with life in the Atlanta of 1985." The attainment of this overall objective involves four consecutive phases of operation as follows [33]: (1) establishing Goals for Education in Atlanta, 1985; (2) developing behavioral objectives that represent the universe of behaviors operationally defining mastery of the Goals at a minimal level necessary for successfully coping with life in the Atlanta of 1985; (3) constructing objective-based, criterion-referenced tests, with test administration procedures, data analysis procedures, and instrumentation, for measuring achievement of the objectives generated in Phase 2; and (4) producing the first round of results from administering the criterion-referenced tests developed in Phase 3. The data base used in this research was collected

in conjunction with phase one. The Delphi technique, a method of soliciting and refining opinion developed by RAND Corporation, was utilized in this phase. This technique was utilized to obtain opinions about the future Atlanta in the light of uncertainty and complexity. In establishing educational goals for the Atlanta of 1985, two types of forecasting are required. One type forecasts what conditions probably will be at a given time in the future, and the other forecasts what educational goals should be in the light of these probable future conditions [3]. In determining what the educational goals in Atlanta should be, the Atlanta Assessment Project utilized position papers written by experts about the future in Georgia. The determination of what the educational goals should be was accomplished through the use of the Delphi technique.

3.1 Delphi Studies

The first Delphi study, conducted in June and July of 1972, involved a panel of 245 professional, technical, managerial, and community leaders in the Atlanta area. The other two studies, conducted simultaneously during November and December 1972, involved a panel of 429 high school teachers, counselors, principals and other administrators directly involved with instruction in the Atlanta Public Schools and a panel of 369 high school student leaders representing all 25 high schools in the Atlanta system. At

each high school, there were 15 students selected to participate in the study.

The initial set of educational goals presented in the Delphi questionnaires was derived from the Goals for Education in Georgia developed by the State Board of Education in 1970. For use in the project studies, the Goals for Education in Georgia were further refined to make the meaning of each goal as concise and understandable as possible and to ensure that each goal statement contained only one substantive element. The revised list contained 86 separate goal statements.

In developing a working model for organizing the goals for education, two considerations were felt to be of importance: (a) reflection of the delphi-developed priorities and areas in the taxonomic structure; and (b) unique assignment of each goal to taxonomic areas. These considerations led to an outline in which the goals were consolidated under five major headings and 16 subheadings which are listed in Table 1. During the remainder of this research the goal areas will be referred to by their associated number designation.

During round one of each study, the respondents gave each of the initial 86 goals an integer rating between one and six. The verbal descriptions associated with each rating are given in Chapter I. Goals 1-86 listed in Table 5 are the initial 86 goals rated during round one. Also during

Table 1. Taxonomic Goal Areas

1.0 Communication Skills

2.0 Personal Development

- 2.1 Self-Understanding
- 2.2 Career Development
- 2.3 Preparation for Life-Long Learning
- 2.4 Preparation for Leisure

3.0 Social Development

- 3.1 Preparation for Citizenship
  - 3.1.1 Commitment to the Principles of Democracy
  - 3.1.2 Forming Relationships with Others
  - 3.1.3 Participating Actively as a Citizen
- 3.2 Preparation for Family Life

4.0 Life Skills

- 4.1 General Goal
- 4.2 Preparation for Managing Health and Environment
- 4.3 Preparation for Managing Personal Finances
- 4.4 Problem Solving

5.0 Academic Skills

- 5.1 Social Sciences
- 5.2 Mathematics
- 5.3 Science and Technology

round one, each respondent was given the opportunity to suggest additional goals that he felt were relevant and should be included.

On the second round of each study, the respondents were asked to rate the original 86 goals plus those additional goals that had been suggested by their group during round one. The respondents were given their group's modal response for round one for each of the original 86 goals during this second round. During round one the community leaders recommended 35 additional goals; the students recommended 16 additional goals; and the educators recommended 14 additional goals. The goals recommended by community leaders, students and educators are listed in Tables 2, 3 and 4, respectively. Thus, on round two the community leaders, students and educators rated 121, 102 and 100 goals, respectively.

Of the additional goals rated during round two of each Delphi study, those goals receiving the highest average ratings were consolidated into a single list. Thirty-five goals were added to the original list of 86 goals. Table 5 lists the original 86 goals and the additional 35 goals which were rated by the students and the educators during round three of their Delphi studies. The community leaders had completed their Delphi study prior to the time that the final listing was prepared. Consequently, the community leaders simply repeated round two of their Delphi study

Table 2. Goals Recommended by Community Leaders

The Individual...
Understands the profit motive and its importance in providing for our basic needs.
Understands that there are individual differences in capacity among persons, even though equal opportunities should be provided.
Desires to do his best consistently so as to help make the world a better place in which to live.
Understands the miracle of man's progress--anthropology and philosophy combined.
Has an opportunity to become more closely acquainted with his peers and his elders through sensitivity group sessions held by community and social organizations.
Understands that he will perceive positive and negative characteristics in most individuals regardless of their socio-economic background.
Respects, rather than tolerates, individuals who are different.
Desires to achieve excellence and develop his own potential to the fullest.
Understands the responsibilities of parenthood and the importance of planning for it.
Is willing to bear arms in defense of this country through the orderly process set forth in our constitution.
Recognizes the applicability to the real world of what he learns through formal education.
Recognizes that specialization may be stifling rather than liberating, in that many specialists work for and are directed by "generalists."
Has the ability to perceive professional or occupational pursuits in the broad context of their relationships to the whole of a given field or the total society.
Appreciates the value of nonprofessional roles in every field of endeavor.

Table 2. (Continued)

The Individual...
Knows how to type proficiently.
Has a working familiarity with current recording and reproducing (storage-retrieval) technology and keeps abreast of changes in this field.
Has a marketable skill upon leaving high school even though planning to pursue further education.
Understands how to initiate legal action for the redress of grievances against individuals or institutions.
Understands what the business world is really like.
Is able to analyze and solve specific problems from an inter-disciplinary point of view.
Understands and appreciates persons of different backgrounds as a result of personal experience with and exposure to many different kinds of persons.
Desires to pursue any given line of inquiry until a rational conclusion or answer is reached.
Has a sense of personal freedom, particularly freedom of expression.
Has a knowledge of cable television and its effects on the society.
Understands and respects the benefits of the free-enterprise system to himself and to the total community.
Has the ability to evaluate the character of other individuals.
Recognizes the difference between believing something because of who says it and believing it because of its true value.
Has the ability to objectively evaluate issues, concepts, problems, etc.
Understands this his interpersonal skills are more important to his success and "coping ability" than are his technical skills.

Table 2. (Concluded)

The Individual...
Understands and appreciates his role in family life.
Is able to recognize a "leading" question.
Recognizes that college is not necessarily for everyone.
Understands the importance of organized, directed group action for dissent or support activity.
Knows how to live in an urban community so as to respect the physical, emotional, and mental well being of others.
Recognizes the value of allowing oneself to dream a little.

Table 3. Goals Recommended by Students

The Individual...
Understands and respects the beliefs and feelings of others
Has a basic knowledge of human psychology, how the human mind and emotions work.
Is informed about opportunities for college acceptance and financial aid to college students.
Knows the basic beliefs of each of the major religions and philosophies of the world.
Knows how to study effectively.
Loves his country.
Is aware of the problems in other countries as well as our own, such as crime, poverty, starvation, etc.
Knows how computers work and their uses.
Knows what he expects in marriage.
Knows and is able to use some skills to save lives (first aid).
Understands how freedom is related to self-discipline.
Knows how and when to give and receive constructive criticism.
Is able to speak one or more foreign languages.
Is able to speak with ease in front of a large group.
Understands the difference between moving up by hard work and ability and moving up by taking things away from other people.
Desires to accomplish something in life that could matter to this world.

Table 4. Goals Recommended by Educators

The Individual...
Is able to ask meaningful questions and to evaluate the answers he gets.
Understands the part played in our society by the mass media--newspapers, magazines, radio, TV, etc.--and the responsibilities of the media.
Understands what is to be gained from having strong family ties.
Recognizes his importance as a member of society.
Understands the difference between "equal" and "identical."
Understands that being educated is not the same thing as having a college degree.
Understands how historical events and patterns are related to present events and patterns.
Understands how mathematics can be used in building motels and tools that may be helpful in providing information for making decisions and solving social, economic, and technical problems.
Is able to put up with tedious, difficult, and sometimes boring activities in working toward a higher goal.
Understands that basic skills in listening, speaking, reading, writing, and mathematics are necessary to be able to do most things of importance in modern society.
Has skills that will make it possible for him to get a job--if he wishes--when he leaves high school.
Understands that success in anything comes only to those who are willing to work for it.
Understands the advantages of attending integrated schools.
Knows how and wishes to create meaningful relationships with others.

Table 5. Goals for Education in Atlanta, 1985--As Seen By Community Leaders, Educators and Students

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
1	values and seeks sound mental and physical health through good nutrition	4.1	C,A	1
2	understands and exercises the citizen's role in the decision-making process of government and politics	3.1.3	C,A	1
3	recognizes that work is necessary and desirable	2.2	C,A	1
4	appreciates the beauty of nature	4.1	A	1
5	knows and practices socially acceptable behavior	3.1.3	C,A	1
6	understands and accepts the necessity and desirability of avoiding discrimination in employment practices	3.1.1	C,A	1
7	understands how technology can alter the natural and physical environment	4.1	C	1
8	is able to function as a follower, a co-worker and a leader in work	2.2	C,A	1
9	understands and accepts the relationship of rights to responsibilities	3.1.1	C,A	1
10	is able to communicate feelings, ideas and information	1.0	C	1
11	understands the judicial system	3.1.2	C	1
12	respects the offices of appointed and elected officials	3.1.3	A	1

See footnote at end of table.

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
13	is able to adjust to changing jobs and job requirements	2.2	C,A	1
14	recognizes and values creativity as a basic human need	2.4	C,A	1
15	understands and accepts the responsibilities and privileges of citizenship	3.1.3	C,A	1
16	has concern for his fellow man	3.1.2	A	1
17	is able to make constructive use of leisure time in some vocational activity	2.4	C,P	1
18	has the ability and desire to participate in community service activities	3.1.3	C,A	1
19	is willing to live in a racially integrated society	3.1.1	C,A	1
20	understands and values the functions, relationships and responsibilities of labor and management in a free society	2.2	C,A	1
21	possesses the attitudes and skills to pursue learning as a life-long process	2.3	C,A	1
22	desires to improve the quality of life in the community	3.1.3	A	1
23	is able to understand and tolerate dissent	3.1.1	C,A	1
24	possesses the knowledge, understanding and appreciation of his heritage	5.1	C,A	1

See footnote at the end of table.

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
25	has a knowledge and understanding of current political issues	5.1	C	1
26	knows how, when and where to secure medical services	4.1	C	1
27	uses one or more of the arts or crafts in recreational and leisure time activities	2.4	C,P	1
28	takes pride in workmanship and accomplishment	2.2	A	1
29	understands and respects himself--his abilities, interests, values, aspirations, limitations	2.1	C,A	1
30	possesses the ability and desire to use the learning resources of the community	2.3	C,A	1
31	possesses a knowledge of and interest in science	5.3	C,A	1
32	is committed to the concept of accountability for the use of public resources	3.1.3	C,A	1
33	knows and understands the concepts of taxation	5.1	C	1
34	is able to make responsible decisions regarding the use of time	4.3	C,A	1
35	recognizes the influence of the family and religious and community organizations in shaping values in a changing society	5.1	C	1
36	understands the structure and functions of local, state and national governments	5.1	C	1

See footnote at the end of table.

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
37	knows and understands workman's compensation, social security, retirement systems, employment insurance and other employee benefits	2.2	C	1
38	values and demands the conservation and proper utilization of land and other natural resources	4.1	C,A	1
39	understands the effects of drugs, alcohol and tobacco	4.1	C	1
40	has the knowledge and skills to be successful in meeting his needs as a consumer of goods and services	4.2	C	1
41	has a knowledge and understanding of international relations	5.1	C	1
42	understands and appreciates the contributions of social, religious and national groups to our culture	5.1	C,A	1
43	understands the social, economic and political implications of population growth	4.1	C	1
44	knows and understands that the quality of man's life depends upon the harmony he achieves with his natural environment	4.1	C	1
45	knows how and where to seek employment and is able to apply for a job and participate in a job interview	2.2	C,A	1
46	possesses the attitudes and personal values that enable him to overcome adversity	2.1	A	1

See footnote at end of table

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
47	has the knowledge and skills for managing personal finances	4.2	C	1
48	has a personal philosophy of life	2.1	C,A	1
49	participates in recreational activities that can provide physical fitness throughout life	2.4	C,P	1
50	understands the techniques to control social and technological change	3.1.3	C	1
51	is able to maintain individual integrity in group relationships	2.1	C,A	1
52	practices responsible behavior when using private and public facilities	3.1.3	C,A	1
53	has knowledge, skills and a desire for life-long growth in arts areas of his choice	2.4	C,P	1
54	understands human biological processes and functions	5.3	C	1
55	has set a tentative occupational or career goal and possesses an educational training plan to achieve it	2.2	C	1
56	understands and is committed to the processes and purposes of law	3.1.1	C,A	1
57	knows how to secure and use community services	4.0.1	C,A	1
58	understands the impact of science and technology on jobs and job requirements	5.3	C	1
59	seeks opportunities to participate in governmental processes	3.1.3	A	1

See footnote at end of table.

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
60	is able to secure information from a wide variety of sources, to analyze, to synthesize, to draw conclusions and to make decisions	4.3	C	1
61	is familiar with a wide variety of occupational fields	2.2	C	1
62	considers public office to be a public trust	3.1.3	C,A	1
63	has an understanding and appreciation of the role of science in our society	5.3	C,A	1
64	is able to identify desirable social and technological changes	3.1.3	C,A	1
65	has the skills necessary for further study or for entry directly into the world of work	2.2	C,A,P	1
66	possesses knowledge and understanding of production, distribution and consumption of agricultrual and industrial products	5.1	C	1
67	knows how and where to obtain additional training and education	2.3	C	1
68	is able to act along or to participate with others in recreational and leisure time activities	2.4	C,P	1
69	understands the functions of public education in our society and how it is administered	5.1	C	1
70	is able to adjust to changing human relationships brought about by geographic and social mobility	2.1	C,A	1

See footnote at end of table

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
71	has the desire to preserve the rights and property of others	3.1.1	A	1
72	is committed to the values expressed in The Bill of Rights	3.1.1	C,A	1
73	desires to acquire and maintain a healthful natural and physical environment	4.1	C,A	1
74	recognizes that every man has the right to participate freely in society so long as the rights of others are not violated	3.1.1	C,A	1
75	understands the emotional and social aspects of human sexuality	3.2	C	1
76	is able to set personal goals	2.1	C,A	1
77	has knowledge of the principle economic, social and political systems of the world	5.1	C	1
78	has knowledge and understanding of mathematics	5.2	C,P	1
79	appreciates the value of the occupations of others	2.2	A	1
80	is able to identify common goals and cooperate with others in their attainment	3.1.3	C,A	1
81	is aware of the social, economic and political implications of technology	5.3	C	1
82	understands freedom as the right to make choices within the framework of concern for the general welfare	3.1.1	C	1

See footnote at end of table

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
83	is able to listen, speak, read and write	1.0	C	1
84	respects and cares for the property of his employer and fellow workers	2.2	A	1
85	understands the capacity of man to adjust to social and technological change	2.1	C	1
86	supports the free and voluntary exercise of religious choice	3.1.1	A	1
87	understands and respects the beliefs and feelings of others	3.1.2	A	3
88	understands and appreciates persons of different backgrounds as a result of personal experience with and exposure to many different kinds of persons	3.1.2	C,A	2
89	knows what he expects in marriage	3.2	C,A	3
90	has a marketable skill (skill that will make it possible for him to get a job, if he wishes) when he leaves high school, even though he may plan to go on for further education	2.2	C,P	2,4
91	has the ability to objectively evaluate issues, concepts, problems, etc.	4.3	C	4
92	understands that being educated is not the same thing as having a college degree	2.2	C	4
93	knows how to live in an urban community so as to respect the physical, emotional, and mental well-being of others	3.1.2	C,A	2

See footnote at end of table

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
94	desires to accomplish something in life that could matter to this world	2.2	A	3
95	understands that basic skills in listening, speaking, reading, writing, and mathematics are necessary to be able to do most things of importance in modern society	1.0	C	4
96	appreciates the value of non-professional jobs in every walk of life	2.2	C,A	2
97	understands that there are individual differences in capacity among persons, even though equal opportunities should be provided	2.2	C	2
98	recognizes that college is not necessary for everyone	2.2	C	2
99	recognizes the usefulness in the real world of what he learns through formal education	2.2	A	2
100	is informed about opportunities for college acceptance and financial aid to college students	2.2	C	3
101	understands the difference between moving up by hard work and ability and moving up by taking things away from other people	2.2	A	3
102	knows how and wishes to create meaningful relationships with others	3.1.2	C,A	4

See footnote at end of table

Table 5. (Continued)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
103	understands the part played in our society by the mass media--newspapers, magazines, radio, TV, etc.--and the responsibilities of the media	5.1	C	4
104	understands how freedom is related to self-discipline	3.1.1	C	3
105	knows how to study effectively	2.3	C	3
106	is willing to bear arms in defense of this country through the orderly process set forth in our Constitution	3.1.3	C,A	2
107	has the ability to evaluate the character of other individuals	3.1.2	C	2
108	knows and is able to use some skills to save lives (first aid)	4.1	C,P	3
109	desires to achieve excellence and develop his own potential to the fullest	2.2	A	2
110	has a basic knowledge of human psychology (how the human mind and emotions work)	5.1	C	3
111	recognizes the difference between believing something because of who says it and believing it because of its true value	4.3	C	2
112	understands the responsibilities of parenthood and the importance of planning for it	3.2	C,A	2
113	understands that his skills in getting along with other people are more important to his success and "coping ability" than are his technical skills	2.2	C	2

See footnote at end of table

Table 5. (Concluded)

No.	Description	Area	[1]Domain	[2]Source
	The Individual...			
114	has a sense of personal freedom, particularly freedom of expression	3.1.1	A	2
115	knows how and when to give and receive constructive criticism	3.1.3	C,A	3
116	is able to ask meaningful questions and to evaluate the answers he gets	4.3	C	4
117	understands how historical events and patterns are related to present events and patterns	5.1	C	4
118	understands what is to be gained from having strong family ties	3.2	C,A	4
119	understands that success in anything comes only to those who are willing to work for it	2.2	C	4
120	is able to speak with ease in front of a large group	1.0	C,A,P	3
121	recognizes his importance as a member of society	2.1	A	4

[1]The letters indicate the domain of the Taxonomy of Educational Objectives relevant to the goal. Thus, C stands for Cognitive Domain; A stands for Affective Domain; and P stands for Psychomotor Domain.

[2]1--indicates goal was one of the original 86.
2--indicates goal was suggested by community leaders.
3--indicates goal was suggested by students.
4--indicates goal was suggested by educators.

during round three. The final column of Table 5 indicates whether the goal stated was one of the original 86 goals or one of the additional 35 goals recommended by community leaders, students, or educators, respectively. Among the 35 additional goals, goals 87 through 121 in Table 5, 15 were recommended by community leaders, 11 were recommended by students and 10 were recommended by educators. These three figures obviously sum to 36, but there were only 35 additional goals. The fact that goal number 90, the Individual Has a Marketable Skill (skill that will make it possible for him to get a job, if he wishes) when he leaves high school, even though he may plan to go on for further education, was recommended by community leaders and educators accounts for this discrepancy.

Upon completion of the three Delphi studies there were approximately 32,000 individual preferences to be analyzed. In addition to these preferences there were 10 items of personal information for each of the 1043 respondents. The items of personal information included age, race, sex, years in Atlanta, educational level, etc. The number of individual preferences together with the personal data provide for a comprehensive data base.

3.2 Goals for Education in Atlanta, 1985

As a consequence of the three Delphi studies, a list of 121 Goals for Education in the Atlanta of 1985 was

established. Table 5 contains a list of these goals, the domain of the Taxonomy of Educational Objectives relevant to each goal listed and the source from which each goal was obtained, i.e., one of the original 86 goals, recommended by community leaders, recommended by students, etc.

The goals for education in Atlanta, 1985, were used to develop behavioral objectives for students. These objectives, derived from the goals, operationally define what a young person should have acquired in the way of knowledge, skills, and values by the time he is old enough to graduate from high school if he is to live successfully in the society of the future here in Atlanta [3].

CHAPTER IV

OBJECTIVES

4.0 Introduction

The overall objective of this research is to determine whether community leaders, students and educators, collectively and individually, differ significantly in their opinions of the future importance of secondary educational goals in the Atlanta of 1985. Establishing particular areas and specific goals in which similarities and dissimilarities of preference exist is of vital importance not only for the future but also for the present. Certainly the identification of differences in the perceived importance of educational goals should aid the Atlanta Board of Education in establishing future curriculum, but it should also serve as a warning signal that the current priorities may not be congruent with current needs. If the preference of those who are implementing the educational process differ from the preference of those who are participating in the educational process, then what can be inferred about the learning atmosphere. Are instructors motivated by having to emphasize subject areas that they feel are only moderately important? Are students motivated to learn subject content of courses they perceive to be of little or no importance? What about

the preferences of those that pay a large share of the education expense, the community leaders? These are moderately successful people who should have a fair opinion about what is required from a secondary education for a person to be a productive, useful member of society. This valuable source of input to the developmental and evaluative process should not be neglected. A comprehensive analysis of the information available in a data base of this nature can provide a basis for change and revision of current priorities in the educational system of Atlanta. Perhaps more importantly in the case of the students is the residual effect of their feeling that their ideas, priorities and aspirations form an integral part of the evaluative process. By examining the relative importance attached to specific goals, insight may be gained into the problem of what type of end product community leaders, students and educators believe the schools should produce.

To determine if community leaders, students and educators differ significantly in their preferences for educational goals, the responses to be used for the comparisons must be determined. After this decision is reached, we must ascertain which goals are to be compared. Once the responses and the goals to be compared have been determined, we are ready to perform the statistical analysis. The hypothesis tests involved will be categorized into the following two categories: (a) category one, which determines

whether the preferences of a category of judges changed over time, i.e., whether the preferences of the community leaders or the students or the educators changed between rounds one and two or between rounds two and three; (b) category two, which determine whether the preferences of community leaders and students, community leaders and educators or students and educators differ significantly. Both category one and category two hypothesis tests are conducted within the goal areas outlined in section 3.1 of Chapter III. By conducting the hypothesis tests for each goal area, we utilize the taxonomic structure inherent in the data base and we reduce the possibility of masking the difference in one or more goal areas that might occur if we conducted the hypothesis tests for all goal areas simultaneously.

4.1 Category One Hypotheses

Since the Delphi technique was utilized to obtain a consensus of opinion within each of the three groups of respondents, (i.e., community leaders, students, and educators) one is tempted to use the standard Delphi approach in analyzing the resulting observations. That is, we determine whether the final round produced the smallest opinion spread as expressed by the interquartile range [9]. Reflecting upon this approach, we see that the location and range of the observations is considered; but, the variation among the observations and the amount of consistency between

successive rounds are not considered. Consistency between successive rounds means that a particular category of judges (i.e., community leaders, students or educators) did not change its preferences for the goals between rounds one and two, between rounds two and three, etc. Chapters V and VI delineate the procedures that will be utilized in this research to determine the variation among the preferences and the consistency in preferences between successive rounds for each category of judges.

Category one hypothesis tests test the null hypothesis that there is no preference difference within a particular category of judges from round K to round K+1. The tests may be expressed as:

$$H_o: w_{i_k} = w_{i_{k+1}} \quad \text{for all i within a goal area}$$

and

$$H_1: w_{i_k} \neq w_{i_{k+1}} \quad \text{for some i within a goal area}$$

where k = the round number and w_i = the Ford weight associated with the i^{th} goal. The definition and calculation of the Ford weight for each goal is discussed in sections 5.1.1 and 5.1.3 of Chapter V. For each category of judges the tests are conducted in each of the 16 goal areas defined in Table 1 of Chapter III. Upon completion of the category one

hypothesis tests, category two hypothesis tests can be considered.

4.2 Category Two Hypotheses

After determining the goal areas and the categories of judges on which category two hypothesis tests will be conducted, the areas in which significant preference differences exist between categories of judges can be identified. The method and rationale for determining the goal areas and categories of judges will be explained in detail in section 6.3.2 of Chapter VI.

Category two hypotheses test the null hypothesis that there are no differences in preference between categories of judges. The tests may be expressed as:

$$H_o: w_{i_j} = w_{i_k}, \text{ for all } i \text{ within a goal area}$$

and

$$H_1: w_{i_j} \neq w_{i_k}, \text{ for some } i \text{ within a goal area}$$

where j = category of judge, k = category of judge ($j \neq k$) and w_i = Ford weight for i^{th} goal. There are three categories of judges and 16 goal areas; however, category two hypothesis tests may not be conducted for all goal areas or for all categories of judges.

Upon completion of the hypothesis tests in category two, the precise area or areas in which significant preference differences exist and their probable effects can be specified.

CHAPTER V

METHODOLOGY

5.0 Introduction

Determining which method of statistical analysis to employ in a research project is perhaps the most important decision that the analyst must make. The validity and ultimate value of the inferences drawn are linked directly to the statistical method utilized. Too often, experimenters and researchers are unaware of the numerous underlying assumptions associated with the statistical tests and sampling distributions that they employ to test their hypotheses. This is partly due to the fact that there is a variety of statistical tests available for testing a particular hypothesis. Most of these tests differ primarily in the underlying assumptions concerning the data collected and the statistical model associated with the test. Invalid inferences are most often the result of nonconformity of the data to the necessary assumptions or the use of an inappropriate statistical test.

As noted in section 2.2 of Chapter II, there are many statistical tests that appear to be applicable to the data of this research. However, upon closer examination it is quite evident that all nonparametric tests are not created

equal. In this context, equal implies that given a set of data, the same inference will not necessarily be drawn when various nonparametric tests are used in the analysis. This observation will be substantiated by an example later in this chapter. Thus, given a set of ordinal data, as is the case of the data base in this research, there are various nonparametric and distribution-free tests available for testing particular hypotheses. The term ordinal means that the goals in one category of importance are not only different from the goals in other importance categories but they also stand in some kind of relation to them, i.e., the relation > (greater than or preferred to) holds for all pairs of categories so that a complete rank ordering of categories arises.

5.1 Ford's Procedure

5.1.1 Background

In 1957 Ford [19] developed a solution procedure for obtaining a single rank order for the problem in which all objects are not ranked by all judges, a very common occurrence in practical problems, i.e., in practical problems of any large magnitude, invariably particular responses from one or more individuals will be missing. Using binary comparisons and a maximum likelihood approach, he was able to show, under mild assumptions, that a solution unique up to a proportionality factor does exist and that his method converged to

that unique optimal solution. The assumption necessary to ensure a solution with his procedure is: "In every possible partition of the objects into two non-empty subsets, some object in the second set has been preferred at least once to some object in the first set."

Thus, given M judges and their preferences for N objects, a win-loss matrix, A, is computed. This win-loss matrix may be represented as

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdot & \cdot & \cdot & a_{1N} \\ a_{21} & 0 & a_{23} & \cdot & \cdot & \cdot & a_{2N} \\ a_{31} & a_{32} & 0 & \cdot & \cdot & \cdot & a_{3N} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{N1} & a_{N2} & a_{N3} & \cdot & \cdot & \cdot & 0 \end{bmatrix} \qquad (1)$$

where a_{ij} represents the number of times object i has been preferred to object j ($a_{ii}=0$). Ford's procedure generates a weight for each object such that the a priori probability of obtaining the observed win-loss matrix is maximized. With these probabilities we may compute the a priori probability of obtaining precisely the matrix of results which we in fact did obtain (i.e., the matrix A), under two assumptions; (a) each comparison that takes place between i and j is independent of all other comparisons, and (b) it is drawn

randomly from a binomial distribution with $p = w_i/(w_i + w_j)$ and with the number of ij comparisons being required to be equal to $a_{ij} + a_{ji}$. More precisely, the function to be maximized is:

$$\Pr(A|w_1,\ldots,w_N) = \prod_{i<j} \binom{a_{ij}+a_{ji}}{a_{ij}} \left(\frac{w_i}{w_i+w_j}\right)^{a_{ij}} \left(\frac{w_j}{w_i+w_j}\right)^{a_{ji}} \qquad (2)$$

However, maximizing this function is equivalent to solving the following problem: In the set $\{w_i>0;\ \Sigma w_i = 1\}$, find values w_i that maximize

$$F_A(\underline{w}) = \prod_{i<j} \left(\frac{w_i}{w_i+w_j}\right)^{a_{ij}} \left(\frac{w_j}{w_i+w_j}\right)^{a_{ji}} . \qquad (3)$$

Taking logarithms and partial derivatives leads to the following set of equations which the w_i's must satisfy at a maximum point:

$$W_i = \sum_{j=1}^{N} \frac{A_{ij}w_i}{w_i+w_j} , \quad i = 1,2,\ldots,N, \qquad (4)$$

where $W_i = \Sigma a_{ij}$, $A_{ij} = a_{ij} + a_{ji}$ and $w_i = \dfrac{\sum_j a_{ij}}{\sum_j a_{ij} + \sum_i a_{ji}}$.

To satisfy this relation and to ensure the existence and uniqueness of a solution, Ford developed an iterative technique

which gives a weight to object i on the k^{th} iteration by the relation

$$w_i^k = \frac{W_i}{\sum_j \frac{a_{ij}+a_{ji}}{w_i^{k-1}+w_j^{k-1}}} \quad . \tag{5}$$

Substituting for W_i we obtain,

$$w_i^k = \frac{\sum_j a_{ij}}{\sum_j \frac{a_{ij}+a_{ji}}{w_i^{k-1}+w_j^{k-1}}} \quad , \tag{6}$$

where w_i^{k-1} = weight assigned to object i on the (k-1)st iteration. The procedure iterates until the weights stabilize. Stabilization occurs when

$$\frac{|w_i^n - w_i^{n-1}|}{w_i^n} < \lambda, \text{ for } i = 1,2,\ldots,N,$$

where λ is selected by the experimenter. For this research λ was selected as .005.

5.1.2 Advantages

Andrews and Pelz [1] have noted that Ford's procedure is most useful when the data have one or more of the following characteristics: "(a) some objects are unclassified by some judges (e.g., when some judges lack knowledge about some of

the objects being ranked), (b) there is less than perfect agreement among the judges, (c) some objects are tied--i.e., placed in the same ranked categories by certain judges, (d) the judges differ in their ability to discriminate along the dimension being considered, (e) the judges differ in their perception of the overall distribution along the dimension being considered (e.g., one judge might feel that highs were rather common, while another might feel that they were rare)." They also observed that if the percentage of wins in the win-loss matrix, A, is used as the initial set of w_i's, this will usually minimize the number of iterations required to obtain convergence of the weights. A variation of their computerized program of Ford's procedure will be used during the analysis portion of this research. A source listing of the program and data input directions are contained in Appendix D.

5.1.3 Application to Data Base

Ford's procedure will be utilized to analyze the data in this research for the following reasons: (a) The preference opinions were collected from each respondent on three separate occasions; and, for various reasons, there are some missing data on one or more rounds for particular individuals; (b) needless to say, there is less than perfect agreement among each of the three groups; (c) since the respondents are categorizing from 86 to 121 goals into one of six levels of perceived importance, numerous goals are tied, i.e., placed

in the same importance category; (d) the respondents differed considerably in their perception of the overall distribution with respect to the importance criterion being used. That is, some respondents felt that the majority of the goals were of the highest importance while others felt that most of the goals were only of average importance. As stated earlier, Andrews and Pelz found that Ford's procedure is most useful under these conditions.

Additionally, this procedure not only provides an excellent overall ranking of the goals; but the relative magnitudes of the weights provide far more insight into how much more one goal is preferred to another. The most commonly used ranking procedures only give an ordinal preference for the goals, which essentially implies that the magnitudes of the preferences for the goals are equally spaced over an integer continuum. Thus, if ranking procedure A supplies us with a ranking of 10 objects, all we can assume is that the objects are on an equal interval continuum. However, Ford's procedure supplies us with a set of weights that is seldom, if ever, on an equal interval continuum. These weights may be interpreted as odds in the sense that the probability that object i will be preferred to object j in a future comparison is $w_i/(w_i+w_j)$. Also, under a null hypothesis of no preference difference between two rating groups, regardless of the number of respondents in each group, the Ford weights will be identical for the two groups.

5.2 Randomization Distribution

Now that the procedure for obtaining observational values has been determined, the statistical test and the associated sampling distribution, known or assumed, must be determined. Of course, the test statistic and its sampling distribution are dependent upon the hypothesis to be tested. This research is concerned with testing the amount of preference difference within goal areas for a category of judges and between all categories of judges. A test statistic which has a sampling distribution derived from a principle first conceived by R. A. Fisher [18] will be used. The application of Fisher's principle is generally referred to as the Method of Randomization. Fisher's principle states that for any set of data we may distinguish between (a) the absolute magnitude of the observations; (b) their algebraic signs; and (c) their locations in the data table denoting the combinations of sampling conditions under which they were observed [7]. Thus, any set of N observations can be considered as one member of a family of distinguishably different sets of N observations having the same absolute magnitudes but **differing in algebraic signs, locations, or** both. By holding the absolute magnitudes constant and considering all combinations of the observations with respect to algebraic sign and magnitude, the exact distribution under the null hypothesis can be determined. This is usually referred to as the Randomization Distribution for the observed

sample. A test statistic is calculated for the sample and for every possible combination of the observations as considered above, thereby obtaining a null distribution for the test statistic. The value of the test statistic for the observed sample is compared with the null distribution and accepted or rejected depending on its relation to the rejection region. The only required assumptions are that the sampling was random, the observations are independent and the test statistic assumes positive and negative values with equal probability. As Bradley [7] has noted, "Thus randomization tests appear to be superior or equal to their parametric counterparts in the generality of cases, superiority giving way to equality in those cases where the parametric assumptions are met." The randomization test makes no distribution assumption nor any assumptions concerning the moments associated with a set of observations. Hence, no assumptions regarding the underlying distribution of the observations nor the interval continuum between the observations are made.

5.3 Comparative Example

To illustrate the methodology developed in the preceding two sections and its advantages over other applicable analysis methods, consider the following example. Suppose there are two groups of judges who each rank 14 objects in order of preference. Let N_1 and N_2 represent the

number of judges in groups one and two, respectively. Applying Ford's procedure to these preferences we obtain a Ford weight for each object ranked by each group. The values listed under groups one and two, respectively, are the Ford weights for each object ranked.

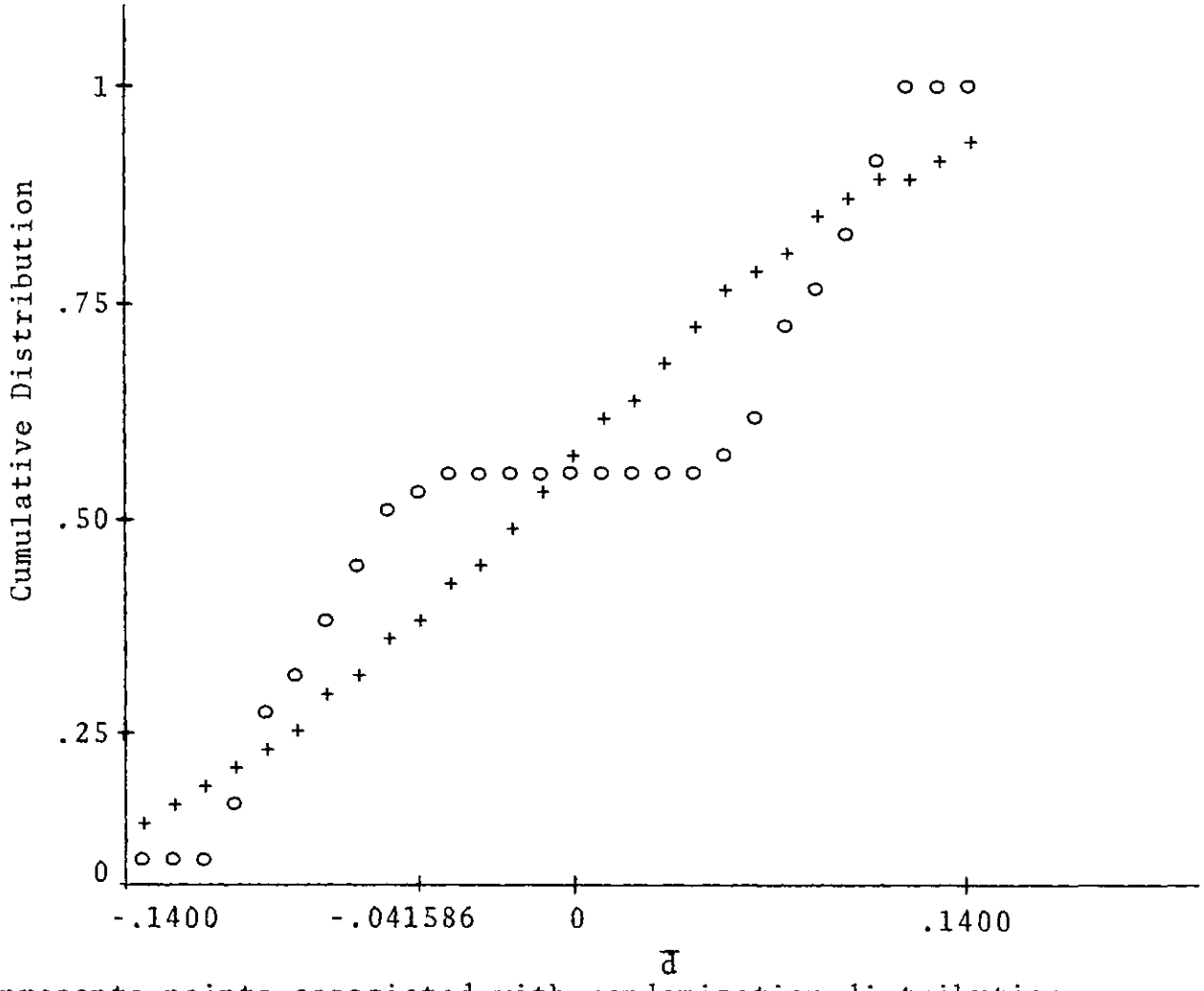
Group 1	Rank	Group 2	Rank	d_i
.686671	1	.446677	4	.239994
.640750	2	1.816666	1	-1.175916
.597647	3	.557700	2	.029947
.521233	4	.479815	3	.041418
.153566	5	.117801	5	.035765
.129318	8	.077970	9	.051348
.142866	6	.107283	6	.035583
.077826	10	.047970	10	.029856
.137462	7	.091802	7	.045660
.049472	12	.037537	11	.011935
.050899	11	.028439	12	.022460
.021380	13	.010102	13	.011278
.011797	14	.006888	14	.004909
.120678	9	.081121	8	.039557

To the right of each Ford weight in each group is the rank of that weight with respect to the other weights within the group. The rank of one is assigned to the largest Ford weight in each group, the rank of two is the next largest, etc. The d_i column is obtained by subtracting the Ford weight for object i in group two from the Ford weight for object i in group one, i.e., letting x_i represent the Ford weight for object i in group one and y_i the Ford weight for object i in group two, we have $d_i = x_i - y_i$. We wish to test the null hypothesis of no preference difference between the two groups of judges. To utilize the randomization distribution we must calculate the value of our test

statistic $\bar{d}$. Since $\bar{d}$ is the average difference observed, we have $\bar{d} = \sum_{i=1}^{14} \frac{d_i}{14} = -.041586$. Under the null hypothesis $\bar{d}$ has an equal probability of being positive or negative. Likewise, each observed difference takes on positive and negative values with equal probability under the null hypothesis. Hence,

$$\bar{d} = \frac{\pm .239994 \pm 1.175916 \ldots \pm .039557}{14} \tag{7}$$

describes the 2^{14} possible sign and magnitude outcomes that $\bar{d}$ can assume. The number of outcomes that are greater in absolute magnitude than the outcome actually observed is 16,236. This value is interpreted as $P_r(|\bar{d}|>.041586) = \frac{16,236}{2^{14}} = .991$. This value is called the p value and will be referred to as such during the remainder of this research. Figure 1 shows the randomization distribution and the t distribution for the observed differences in this problem. The two distributions and the p value were obtained via a computer program executed on the Univac 1108 computer. To avoid a possible misinterpretation of the distributions it should be pointed out that the randomization distribution is scaled according to $\bar{d}$ and the t distribution is scaled according to $\frac{\bar{d}\sqrt{n}}{s}$, where n is the sample size and s is the sample standard deviation. These comments also pertain to Figures 3, 4, and 5.



o represents points associated with randomization distribution
+ represents points associated with t distribution

Figure 1. Example Problem: Randomization Versus t Distribution

Since the t distribution is calculated for these observations, let's examine the probability level associated with the t statistic. Recall that the t test is a parametric test and requires the assumption that the variable in question is from a normal distribution. For this particular sample we find that the test statistic t_o is significant at the .608 level, i.e., the probability of the test statistic t_o being greater than the observed value is .608.

To compute the Kruskal-Wallis test statistic for the two groups, we must obtain a single rank ordering of all 28 Ford weights. Listed below is the overall rank for each Ford weight in each of the two groups.

Group 1	Group 2
2	8
3	1
4	5
6	7
9	14
12	18
10	15
19	22
11	16
21	23
20	24
25	27
26	28
13	17
Rank Sum = 181	Rank Sum = 225

As noted in section 2.2.2 of Chapter II, the Kruskal-Wallis test statistic H is given by

$$H = \left(\frac{12}{n(n+1)}\right)\left(\sum_{j=1}^{k} \frac{T_j^2}{n_j}\right) - 3(n+1) \qquad (8)$$

where n = total number of observations, n_j = number of observations in the j^{th} group and t_j = the rank sum of the j^{th} group. Hence,

$$H = \left(\frac{12}{28.29}\right)\left(\frac{(181)^2}{14} + \frac{(225)^2}{14}\right) - 3(29) = 1.02 \qquad (9)$$

and from the χ^2 table with one degree of freedom the probability of $H \geq 1.02$ is approximately .33 when the null hypothesis is true.

To use Wilcoxon's test for paired observations we must compute the rank differences, the sum of the negative ranks and the sum of the positive ranks as shown below:

Group 1	Group 2	Group 1-Group 2	Ranks	Ranks(+)	Ranks(-)
1	4	-3	8		8
2	1	1	3-5/7	3-5/7	
3	2	1	3-5/7	3-5/7	
4	3	1	3-5/7	3-5/7	
5	5	0	--		
8	9	-1	3-5/7		3-5/7
6	6	0	--		
10	10	0	--		
7	7	0	--		
12	11	1	3-5/7	3-5/7	
11	12	-1	3-5/7		3-5/7
13	13	0	--		
14	14	0	--		
9	8	1	3-5/7	3-5/7	
				Σ=18-4/7	Σ=15-3/7

After obtaining the rank difference of group one minus group two, we eliminate all zero differences from further consideration. Next we rank absolute rank differences beginning with the smallest value. The ranks associated with positive rank differences are listed as are the ranks associated with negative rank differences. Let the sum of the positive rank values be T_1 and the sum of the negative rank values be T_2. As noted in section 2.2.2 of Chapter II, if $n \geq 8$ then the distribution of T_1 and T_2 is approximately normal. Thus

$$z = \frac{T_1 - \mu_T}{\sigma_T} \tag{10}$$

and

$$z = \frac{T_2 - \mu_T}{\sigma_T} \tag{11}$$

where

$$\mu_T = \frac{n(n+1)}{4} , \tag{12}$$

$$\sigma_T = \sqrt{\frac{(2n+1)\mu_T}{6}} \tag{13}$$

and z has a distribution that is approximately that of a standard normal distribution. In these two expressions, n is the number of ranks remaining in the rank difference column

after the zero entries are eliminated. Hence,

$$\mu_T = \frac{n(n+1)}{4} = \frac{8.9}{4} = 18 \qquad (14)$$

and

$$\sigma_T = \sqrt{\frac{(2n+1)\mu_T}{6}} = \sqrt{\frac{(17)18}{6}} = 7.1 \ . \qquad (15)$$

Thus,

$$z = \frac{18-4/7-18}{7.1} = .08 \text{ and } z = \frac{15-3/7-18}{7.1} = -.08 \ . \qquad (16)$$

For a two-tailed test we find from the standard normal table that the probability that $z \geq .08$ or $z \leq -.08$ is approximately .936.

In summary, the inference that is drawn concerning the null hypothesis that the two samples are from a common population is dependent upon the statistical test utilized. The method proposed in this research yielded a p value of .991 and the Wilcoxon test provided a value of .936. However, the Kruskal-Wallis test yielded a value of .33 which is considerably lower than the probability from the other tests.

The t distribution gave a value of .608, and from Figure 1 we can see that had d been slightly less, say -.05, the significance level from the t distribution would have been much more different than the actual level. The actual

level is that one produced by the randomization distribution for the observed differences. Also, as noted earlier, the use of the t distribution requires the assumption of normality, which, as Figure 1 shows, is obviously unwarranted for this data. Conversely, the randomization distribution requires no assumptions regarding any of the moments of the underlying distribution.

CHAPTER VI

DELPHI TECHNIQUE--CATEGORY ONE HYPOTHESIS TESTS

6.0 Introduction

Consensus research is the study of techniques to determine the most reliable consensus if several expert opinions are offered for a particular estimate [9]. The Delphi technique, a prominently used consensus technique, was originally developed by the RAND Corporation for obtaining greater consensus of opinion among experts about urgent defense problems without face-to-face discussion [50]. As Kelly and Thibaut [27] have noted, in direct confrontation, group opinion is highly influenced by dominant individuals and the results are often distorted because the weaker individuals tend to conform to group pressure [2].

The objective of the Delphi technique is to induce opinion convergence through a sequence of questionnaires interspersed with controlled opinion feedback [9]. The technique, when employed properly, allows for independent thought and evaluation on the part of each respondent while assisting him in thinking through his opinions by providing controlled feedback information. When dealing with a group of equally competent individuals, the elicitation and refinement of opinions in this manner are essential, according to Dalkey [11].

He noted that divergence of answers is more likely to occur when we are dealing with an individual's opinion as opposed to more solid knowledge. RAND is in the process of conducting experiments to determine if Delphi works as well in evaluating goals as it does in areas of factual judgment. The preliminary results, according to RAND's current Delphi expert, Norman C. Dalkey [26], are "Beautiful, better than we hoped."

6.1 Application of the Delphi Technique in Studies of Educational Goals

Several applications of the Delphi technique in the area of educational goals have surfaced in recent years. Cyphert and Gant [10] used the Delphi technique to identify prime targets on which the School of Education at the University of Virginia should concentrate its energies and resources for the next decade. The participants included: (1) faculty members, graduates and undergraduates from the School of Education; (2) deans, the president's cabinet and elected members of the university faculty senate; (3) off-campus educators, i.e., school teachers and administrators holding elective office in statewide professional organizations and the deans of the State's major schools of education; (4) officers of the Virginia School Boards Association, the Virginia PTA and the State Council of Higher Education; (5) political leaders, e.g., the education committees of the

State Legislature, U. S. Senators and Representatives, the Governor, etc.; and (6) leading newspaper editors and persons in the Virginia AFL-CIO, NAACP, and Chambers of Commerce involved in education. During phase one more than 750 suggestions were reduced to 61 generic statements. The respondents then rated the statements on a five-point priority continuum. On questionnaires III and IV, the respondents were given the modal response and their previous responses; on the latter they were also provided with a list of the major "dissenting opinions." The authors noted that their application of the Delphi technique differed from previous applications in the following four ways: (1) groups of 400 respondents were used as opposed to the 50 or fewer used in previous studies; (2) participants were not all experts in the education field; (3) the technique was used to predict what should happen as opposed to what will happen; and (4) the mode, rather than the interquartile range, was used as a measure of consensus. Their study showed that the technique can be used to mold as well as to elicit the opinions of respondents and that the greatest amount of change in opinion occurs on the first questionnaire in which the modal response is supplied.

Utilizing the Delphi technique, Anderson [50] assisted a county school district in clarifying and setting objectives; and, Norton [50] assisted in identifying the needs for a newly planned university. Judd [26] noted that

the possibilities for the Delphi technique's usefulness extend far beyond those of long-range planning into the realm of expedient, low-cost and democratic consensus formation on today's problems. He also cited an application of the technique by a liberal arts college's committee charged with developing the curriculum for a new branch campus. In a matter of days a highly innovative and experimental type of curriculum was adopted by a conservative faculty.

Uhl [50] employed the Delphi technique to assess the present and preferred importance of goals of five colleges and universities with quite different characteristics. He used parametric methods and employed a repeated measures design and multivariate analysis of variance (MANOVA) to investigate convergence of opinion among the respondents. This procedure uses the absolute sum of the differences between each participant's response and the mean of all participants' responses. Although he did not use respondents who could be considered experts on educational goals, he concluded that convergence did not lead to less accurate data.

Education, as one of the most important institutions of our society, is caught up in this modern world of rapid change and uncertain future. These applications of the Delphi technique indicate that it can be an effective tool in the determination and evaluation of future educational requirements.

The actual procedures employed in the studies above vary in one or more aspects from the general Delphi procedure.

The researcher should be aware of the general Delphi procedure and the effects that any changes to that procedure may have on the final results.

6.2 Procedure

6.2.1 General Procedure

There are numerous variations in the actual application of the Delphi technique. The general procedure is as follows:

(1) The opinions of a group of experts are solicited on a specific topic.

(2) The respondents are asked to evaluate or rank the list of opinions by a criterion such as feasibility, importance, probability of occurrence, etc.

(3) The respondent is again presented with the list and supplied appropriate feedback concerning the previous responses. If his response is outside a specific range, he may be asked to revise his opinion or state the rationale for his position. Also, the respondent may be supplied with his previous response or responses.

(4) Step three (3) is usually repeated, and the procedure terminates.

6.2.2 Procedure Utilized in Collection of Data Base

The data base for this research, as described earlier, consists of three separate and identically conducted Delphi studies. These studies differed from the standard Delphi

technique in the following four ways:

(1) The respondents were not necessarily experts per se in the area of educational goals.

(2) The respondents were not required to provide the initial list of goals to be considered in subsequent phases of the study.

(3) The respondents were required to state their reasons for being outside a specific range on questionnaire two only.

(4) The respondents in the student and educator Delphi studies were not supplied with their previous responses at any time.

Although the respondents were not experts per se, it was felt that each group possessed a particular input that was desired in the overall evaluative effort. Experiments by Brown, Cochran and Dalkey, using students as participants, suggest that there is no great loss in including less knowledgeable individuals as long as some of the participants are knowledgeable about the subject area.

The respondents were not required to submit the initial list of goals because it was felt that the Goals for Education in Georgia, developed by a panel of educational scholars, would provide an adequate starting point. The respondents were allowed to submit goals to be added to the initial list. It should be noted that in the original list of goals, none received an average rating of one, i.e., on

the average, none of the original 86 goals were rated as being of little or no importance.

The respondents were required to justify their responses on questionnaire two only because of time constraints, the large number of respondents and the manpower allocated for the study. The respondents in the student and educator studies were not supplied with their previous responses for the reasons just mentioned. It was believed that omitting these steps would not substantially affect the benefits obtained from the Delphi technique.

6.3 Determination of Convergence

6.3.1 Methods Normally Employed

In attempting to ascertain whether the Delphi technique has been successful in obtaining a convergence of opinion, one of three criteria is normally employed. The responses of the individuals are compared with the modal response or with the mean response in some manner or the number of individual responses that lie within the interquartile range is counted. Using the mean or mode and the responses from a category of judges between successive rounds to determine the amount of convergence requires, as a minimum, an assumption about the first moment of the underlying distribution. Usually an assumption regarding the second moment of the underlying distribution is also required. As we illustrated in section 5.3 of Chapter V, normality

assumptions are not always reasonable. Regarding the criterion of the number of responses being within the interquartile range, no statistical test is usually applied other than noting the number of responses that fall within the range from questionnaire to questionnaire. Frequently, it is also noted that the interquartile range becomes smaller with succeeding questionnaires. Another criterion that has been suggested is that of standard deviation reduction. One advantage of this criterion is that it is testable using nonparametric statistical tests. It suffices to say that these methods of analysis are limited by their degree of resolution of the actual convergence and by their underlying assumptions.

6.3.2 Recommended Method

The method of analysis proposed in this research is derived from Fisher's Principle of Randomization. It requires no assumptions about any moments belonging to the underlying distribution. The method involves the use of the randomization distribution, the exact distribution for the observed data, which possesses all of its moments. Through the use of this distribution and an appropriate test, an exact statistical inference concerning the amount of consistency between rounds within each category of judges is obtained. This statistical inference is independent of the underlying distribution of each judge's responses and of the entire group of judges' responses.

As previously noted, there were three Delphi studies

conducted in conjunction with the Atlanta Assessment Project. Each of the 1043 respondents was asked to rate a list of educational goals on their perceived importance utilizing an integer continuum of one to six. For each questionnaire of each Delphi study, Ford's procedure is applied to obtain a weight for each goal rated. These weights may be used directly to obtain a preference ranking of the goals for that questionnaire. Although the transformation of weights to ranks is useful in itself, much information concerning the relative magnitudes of the weights is lost by their conversion. By allowing the goals to remain associated with the actual weights generated by Ford's procedure, we have a clearer indication of the degree of importance associated with each goal for a particular questionnaire of the Delphi study. If the weights are not converted to ranks, what statistical test can be used to make valid inferences? Before answering this question, it is necessary to ascertain the statistical hypothesis that is to be tested. The succinct question is, "Has the Delphi technique been successful in obtaining a consensus of opinion?" To resolve this question, a criterion for measuring consensus must be determined. The criterion most often used is the amount of convergence that occurs on successive questionnaires. In this research consensus will be measured by the degree of consistency in a preference observed between successive questionnaires for a category of judges.

The methodology outlined in Chapter V was used to ascertain the degree of consensus obtained by the Delphi technique. Given the importance ratings from two successive questionnaires for a particular category of judges, the null hypothesis that there are no preference differences between the judges on the two questionnaires was tested. Ford's procedure was utilized to generate a weight for each goal rated in the two questionnaires. A set of differences was obtained by subtracting the weight for each goal in one questionnaire from the weight of the corresponding goal in the next questionnaire. The average of these differences is our test statistic $\bar{d}$. Using the randomization distribution associated with the observed sample and our test statistic, we can make an exact probability assessment concerning the differences in the Ford weights from the two questionnaires. If the probability assessment does not indicate that the null hypothesis should be rejected, no significant changes in opinion were observed between the two questionnaires. When no significant differences in preference are observed between questionnaires, we can interpret this to mean that the preferences of the judges within a category are consistent. Theoretically, when consistency is obtained, the weights will be identical in each succeeding questionnaire. The graph in Figure 2 represents a typical set of values obtained from using the randomization distribution to measure consistency within a category of judges between questionnaires. The k^{th}

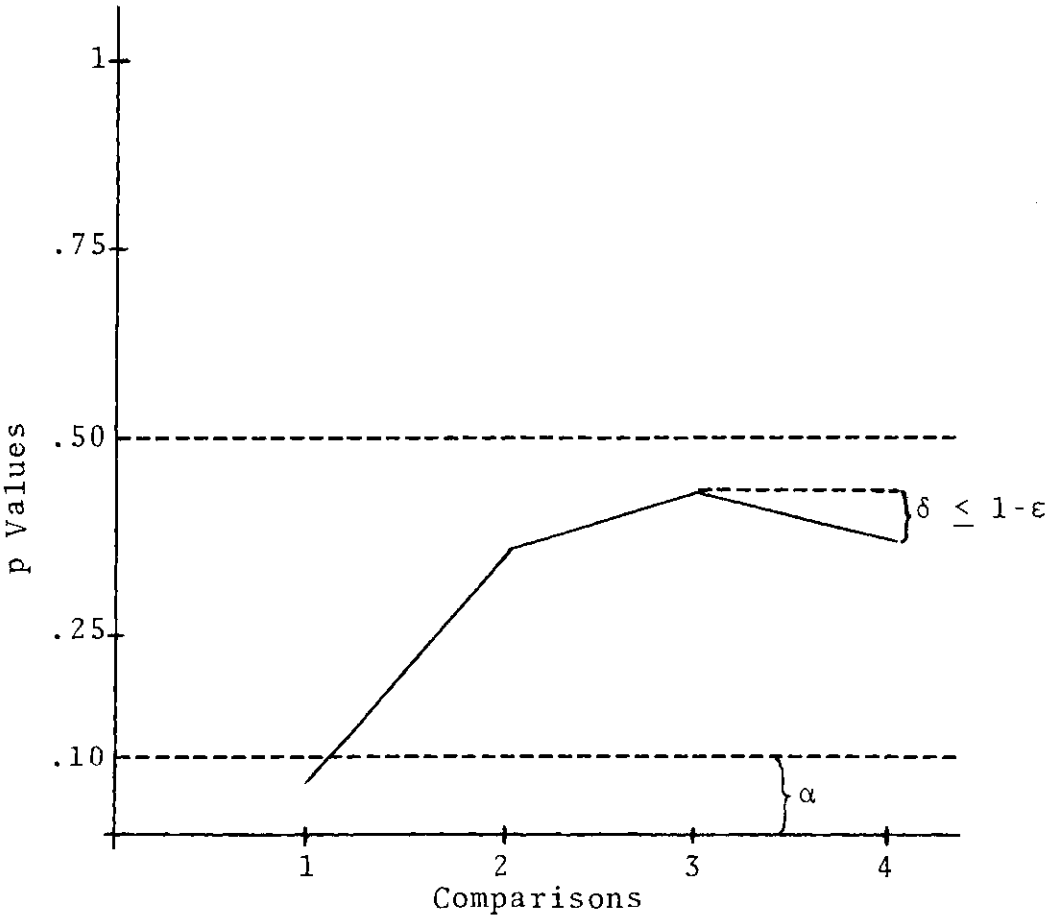


Figure 2. Consistency in Preference Graph

questionnaire comparison refers to a comparison of the weights between the k^{th} and $(k+1)^{st}$ questionnaires in which a category of judges rated the goals. The values plotted are the significance levels at which the null hypothesis of no preference difference could be rejected, i.e., the p values associated with the two questionnaires being compared. The portion of the curve from one to three comparisons can be thought of as the "learning or rethinking" curve. During this period the respondents are evaluating their previous ratings with the feedback information and rethinking each question on the questionnaire.

Viewing the Delphi questionnaire in this manner allows the experimenter to determine when the desired level of consistency has been attained. This method also allows the experimenter to determine the time required to reach consistency for the type of feedback information provided during the Delphi study. Additionally, the experimenter can delete those subsections in which consistency has occurred from the succeeding questionnaires. The method also enables the experimenter to know how much value to place on the information obtained from the Delphi study. If the p value from the randomization distribution is less than α, then the value of the information contained in the questionnaires is highly questionable. The null hypothesis of no differences in preference is rejected, i.e., this implies that the preferences were not consistent between questionnaires. Such

information should not be compared to information contained on questionnaires receiving a high p value, i.e., consistent preferences between questionnaires. Perhaps most importantly, this method enables the experimenter to answer the question posed earlier, "Has the Delphi technique been successful in obtaining a consensus of opinion?" By selecting an appropriate ε (epsilon) prior to the study, the experimenter can determine whether the Delphi technique was successful in obtaining consistent preferences by comparing the absolute difference between the last two comparisons with $1 - \varepsilon$, given that the last p value was greater than α.

If the p value associated with the last comparison is less than α, then there is too much variability in the judges' preferences between questionnaires for the information in the final questionnaire to be of much value to the experimenter. However, if the p value associated with the last comparison is greater than α, then the null hypothesis of no differences in preference is not rejected and the question of consistency is addressed. The absolute difference in the p values for the last two comparisons, say δ, is compared to $1 - \varepsilon$. If this absolute difference is less than or equal to $1 - \varepsilon$, then we can say that this category of judges' preferences has converged to consistency.

The selection of ε is crucial to the overall evaluation of the degree of success of the Delphi technique. Suppose we obtained a p value of .05 for the first comparison and a

p value of .95 for the second comparison. Although it is clear that the questionnaires which produced the p value of .95 are in high agreement and possess low variability, we would reject the data as having converged to consistency if $1 - \varepsilon$ is less than .90. Thus, the data contained in the final two questionnaires will also be said to have converged to consistency if the p value associated with the final questionnaire is greater than .50.

To summarize, the steps involved in the determination of whether a category of judges' preferences has converged to consistency are listed below:

1. Calculate a p value for each pair of successive questionnaires.

2. Compare the final p value calculated with the predetermined α level. If the p value is less than α, then that category of judges' preferences contains too much variation to be considered as consistent. The null hypothesis of no differences in preference for that category of judges is rejected. Hence, the value of the data contained in the final questionnaire is suspect and should not be used in any subsequent statistical test between categories of judges.

3. If the final p value is greater than .50, the null hypothesis is not rejected and we can say that the data in the final questionnaire is consistent with the data in the previous questionnaire. Hence, the Delphi technique has succeeded in producing consistent preferences.

4. If the final p value is greater than or equal to α but less than .50, we look at the absolute difference of the last two p values, say δ. If this absolute difference is less than or equal to $1 - \varepsilon$, the preferences in the questionnaires are consistent.

6.4 Comparative Example

To illustrate the proposed methodology, consider the following example. Given the preferences of the students in the three rounds of their Delphi study for goal area 2.2, "Career Development," we desire to determine whether the Delphi technique has produced consistency within this category of judges' preferences. The preferences indicated on each of the three questionnaires are used to generate a Ford weight for each of the goals. The Ford weights generated for each questionnaire are listed in Table 6. Also listed in Table 6 are the differences of the Ford weights for questionnaires two and three. To determine the consistency in preference obtained between questionnaires one and two, the test statistic $\bar{d}$ is calculated to test the null hypothesis of no difference in preferences between the two questionnaires. For questionnaires one and two the value of $\bar{d}$ is -.707199. Using a computer program designed to generate the null distribution of $\bar{d}$ and the t distribution associated with $\frac{\bar{d}\sqrt{n}}{s}$, the distributions shown in Figure 3 are those derived from the results of questionnaires one and two. The p value

Table 6. Goal Area 2.2: Student Delphi Study

Goal Number	Round One	Round Two	Round Three	Round One Minus Round Two	Round Two Minus Round Three
3	.629599	1.303263	.696169	-.673664	.607094
8	1.025170	2.395940	.976808	-1.370770	1.419132
13	.414352	.092247	.176578	.322105	-.084331
20	.247868	.023427	.042432	.224441	-.019005
28	.553458	1.651736	.965433	-1.098278	.686303
37	.572545	1.588100	1.195374	-1.015555	.392726
45	.985040	3.129809	1.568098	-2.144769	1.561711
55	1.108244	1.980788	1.540770	-.872544	.440018
61	.221096	.028022	.042601	.193074	-.014579
65	1.624591	4.180566	2.704901	-2.555975	1.475665
79	.150774	.020703	.039845	.130071	-.019143
84	.469132	.093657	.222435	.375475	-.128778
90	*	*	*	-	-
92	*	*	*	-	-
94	*	.426919	.884819	-	-.457900
96	*	*	*	-	-
97	*	*	*	-	-
98	*	*	*	-	-
99	*	*	*	-	-
100	*	.277658	1.076102	-	-.798444
101	*	.300718	2.226362	-	-1.925644
109	*	*	*	-	-
113	*	*	*	-	-
119	*	*	*	-	-

*Indicates that the goal was not rated on that round of the study.

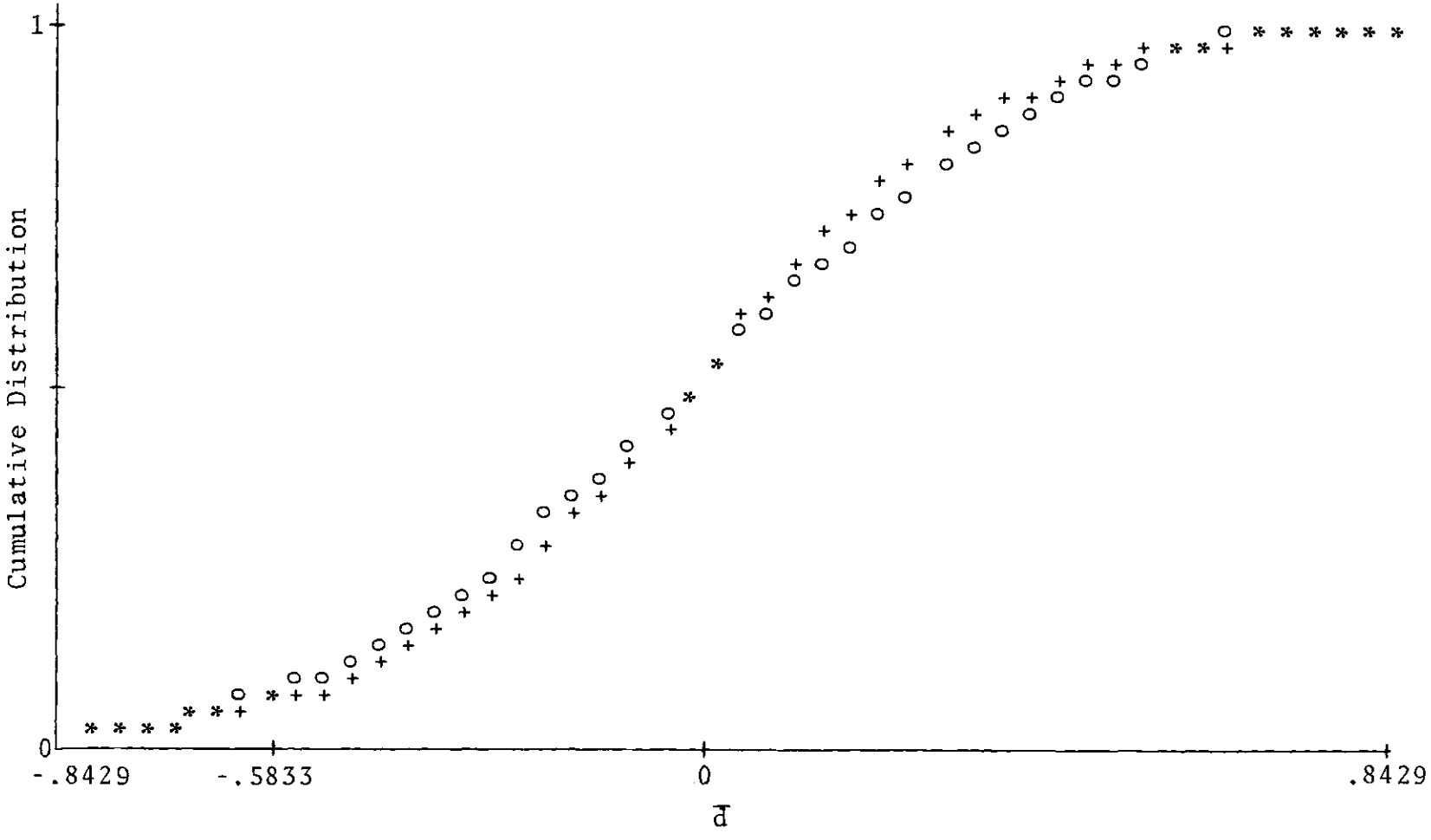
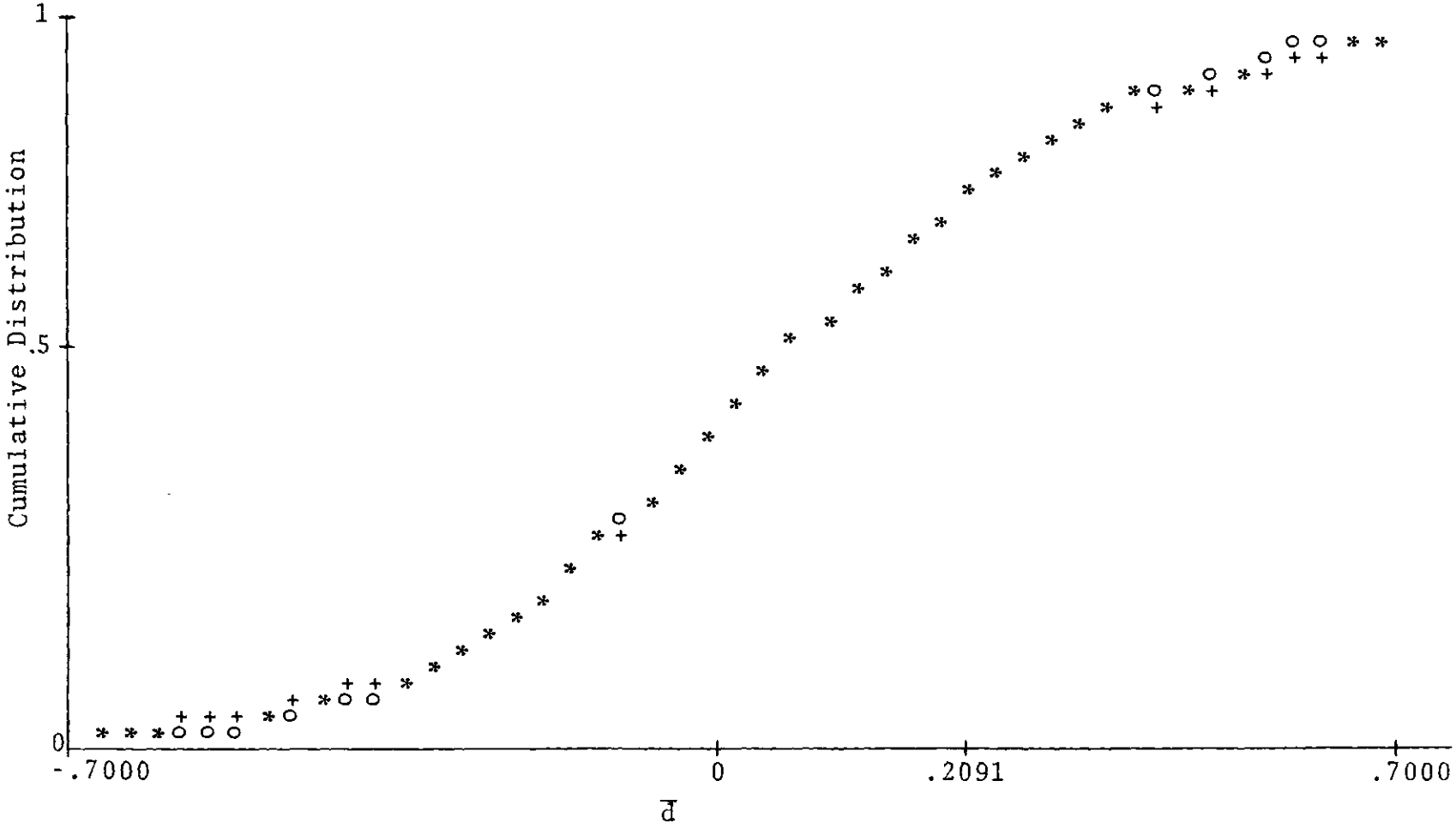


Figure 3. Career Development: Student Delphi Study, Rounds One and Two

yielded by the randomization distribution is .032. Assuming an α level of .10, we would reject the null hypothesis of no difference in preferences between questionnaires one and two. To test for consistency in preference between questionnaires two and three we calculate $\bar{d}$ and find that it is .208988. Figure 4 shows the null distribution for $\bar{d}$ and the t distribution associated with $\frac{\bar{d}\sqrt{n}}{s}$. The p value for questionnaires two and three is .426. Using an α level of .10, we would fail to reject the null hypothesis of no difference in preferences between questionnaires two and three. Since the observed p value is less than .50 but greater than α, the absolute difference in the two p values obtained for this problem must be calculated and compared with $1 - \varepsilon$. Letting $\varepsilon = .50$, we see that the absolute difference is less than the value of $1 - \varepsilon$, i.e., .394 is less than .50. Hence, the students' preferences are consistent in goal area 2.2, "Career Development."

The t distribution, as shown in Figures 3 and 4, gives a good approximation to the exact distribution of the data for questionnaires one and two and questionnaires two and three. The probability of obtaining a larger value for the t test statistic than was observed is .037 for questionnaires one and two and .415 for questionnaires two and three. For this goal area and this category of judges, the assumption of normality appears tenable.



o represents points associated with randomization distribution
+ represents points associated with t distribution

Figure 4. Career Development: Student Delphi Study, Rounds Two and Three

Employing the criterion of standard deviation reduction, we rank the Ford weights for questionnaires one and two and compute the rank differences required for the Wilcoxon test. As explained in section 2.2.2 the required rankings are:

Round 1	Round 2	(Round 1-Round 2)	Rank	Rank (+)	Rank (-)
5	3	2	7	7	
11	10	1	3	3	
6	4	2	7	7	
8	7	1	3	3	
1	6	-5	$9\frac{1}{2}$		$9\frac{1}{2}$
3	8	-5	$9\frac{1}{2}$		$9\frac{1}{2}$
10	11	-1	3		3
7	9	-2	7		7
2	1	1	3	3	
12	12	0	-		3
4	5	-1	3		3
9	2	7	11	11	
				$\Sigma=34$	$\Sigma=32$

Thus,

$$n = 11,\ \mu_T = \frac{n(n+1)}{4} = \frac{11.12}{4} = 33, \tag{1}$$

$$\sigma_T = \sqrt{\frac{(2n+1)\mu_T}{6}} = \sqrt{1265} = 35.57 \tag{2}$$

and

$$z = \frac{T_1 - \mu_T}{\sigma_T} = \frac{34-33}{35.57} = .028. \qquad (3)$$

From the standard normal distribution tables we find that the probability that $z \leq -.028$ or $z \geq .028$ is approximately .978. Hence, the null hypothesis is not rejected and we conclude that there was no preference difference between questionnaires one and two. For questionnaires two and three the ranks associated with the Ford weights along with the rank differences, positive ranks and negative ranks are listed below:

Round 2	Round 3	Round 2-Round 3	Rank	Rank (+)	Rank (-)
6	1	5	8	8	
13	7	6	9	9	
7	8	-1	$2\frac{1}{2}$		$2\frac{1}{2}$
10	10	Ø	-		
9	5	4	7	7	
11	4	7	$10\frac{1}{2}$	$10\frac{1}{2}$	
14	13	1	$2\frac{1}{2}$	$2\frac{1}{2}$	
12	9	3	6	6	
4	3	1	$2\frac{1}{2}$	$2\frac{1}{2}$	
15	15	Ø	-		
8	6	2	5	5	
5	12	-7	$10\frac{1}{2}$		$10\frac{1}{2}$
3	2	1	$2\frac{1}{2}$	$2\frac{1}{2}$	
1	11	-10	12		12
2	14	-12	13		

Thus,

$$n = 13,\ \mu_T = \frac{n(n+1)}{4} = \frac{13.14}{4} = 45.5, \tag{4}$$

$$\sigma_T = \sqrt{\frac{(2n+1)}{6}\,T} = \sqrt{\frac{27.45.5}{6}} = 14.31 \tag{5}$$

and

$$z = \frac{53-45.5}{14.31} = .524. \tag{6}$$

From the standard normal distribution table we find that the probability that $z \leq -.524$ or $z \geq .524$ is approximately .60. Hence, for questionnaires two and three the Wilcoxon test again fails to reject the null hypothesis of no preference difference. From the results of the Wilcoxon test we infer that there was consistency in preference for all three questionnaires.

Looking at the absolute sum of mean minus individual responses, we find that on questionnaire one the sum was 4016.15 and on questionnaires two and three it was 1629.78 and 2704.99, respectively. Clearly, there was a large reduction between questionnaires one and two; however, there was an increase between questionnaires two and three. Again, it is extremely difficult to ascertain the amount of consistency in preference.

6.5 Results of Category One Hypothesis Tests

The proposed methodology outlined earlier in this chapter for determining whether a category of judges' preferences are consistent was utilized to conduct the category one hypothesis tests. A total of 42 hypothesis tests was conducted within goal areas between questionnaires two and three. Table 7 contains the p values obtained from the randomization distribution for each of the category one hypothesis tests. An α value of .10 and an ε value of .5 were used to determine if consistency in preference was obtained between questionnaires two and three within each category of judges.

Consistency in preference occurred in ten goal areas in the community leader and student Delphi studies and in 13 goal areas in the educator Delphi study. In one goal area, Forming Relationships with Others, of the community leader Delphi study, there was considerably less consistency between rounds two and three than between rounds one and two. However, since the absolute difference in p values is less than .5, a null hypothesis of consistency in preference is not rejected. Employing this argument for two goal areas, Participating Actively as a Citizen and Preparation for Managing Personal Finances, in the educator Delphi study, the null hypothesis of consistency in preference is also not rejected. In the goal area of Preparation for Family Life, we were not able to compare rounds one and two, since the

Table 7. Results of Category One Hypothesis Tests

Pr ($|\bar{d}| \geq$ observed value) = p. (The figures represent values of p.)

Goal Areas	Community Leaders		Students		Educators	
	Rds 1&2	Rds 2&3	Rds 1&2	Rds 2&3	Rds 1&2	Rds 2&3
Communication Skills	.250	.562	.033	.499	.125	.393
Self-Understanding	.466	.859	.119	.092	.047	.629
Career Development	.228	.420	.032	.426	.103	.333
Preparation for Life-Long Learning	.250	.926	.838	.691	.323	.401
Preparation for Leisure	.021	.082	.021	.793	.026	.990
Commitment to the Principles of Democracy	.018	.281.	.048	.431	.002	.236
Forming Relatsionships with Others	.619	.302	.577	.625	.735	.664
Participating Actively as a Citizen	.933	.853	.043	.075	.420	.104
Preparation for Family Life	*	.450	*	.634	*	.50
Life Skills	*	*	*	*	*	*
Preparation for Managing Health and Environment	.002	.004	.023	.068	.056	.102
Preparation for Managing Personal Finances	.279	.373	.519	.942	.522	.208
Problem Solving	.086	.926	.364	.250	.653	.804
Social Sciences	.001	.001	.028	.576	.009	.588
Mathematics	*	*	*	*	*	*
Science and Technology	.043	.059	.045	.003	.063	.063

*Indicates that a statistical test was not possible since there was only one goal in that goal area.

goals in this area were not rated on questionnaire one. The values obtained from comparing questionnaires two and three of this goal area are above .40 for all three Delphi studies. Even though we only have one comparison, we can consider this goal area as being consistent since the values are relatively high. In each of the two goal areas, Life Skills and Mathematics, there is only one goal, so no statistical test was possible.

6.6 Summary

In summary, the proposed method of analysis allows the experimenter to systematically delete subsections from the Delphi questionnaire as the desired level of consistency in preference is reached. It also provides for an evaluation of the feedback information being provided to the category of judges being examined. Most importantly, it provides an exact probabilistic assessment of the consistency in preference observed between questionnaires. The experimenter is then better able to assess the value of the information within each goal area obtained from the Delphi technique. If consistency in preference is not attained in a sufficient number of goal areas, this would indicate to the experimenter that he should: (a) restructure his instrument, i.e., a higher/lower resolution instrument is needed to bring about consistency; (b) alter the type or the amount of feedback information to the judges. By utilizing this method he is

much better prepared to answer the question posed earlier concerning the degree of success achieved by the Delphi technique. He is also better able to identify the specific area or areas that require further investigation.

CHAPTER VII

CATEGORY TWO HYPOTHESIS TESTS

7.0 Introduction

Using the results from the hypothesis tests in category one, we are able to determine the goal areas in which it would be meaningful to perform category two hypothesis tests. Recall that category two hypothesis tests are those that test the null hypothesis of no differences in preference between community leaders and students, community leaders and educators and students and educators, i.e., no differences in preference between categories of judges. That is, we would like to know in which specific goal areas those outside the academic environment differ from those inside the academic environment and in which goal areas those inside the academic environment differ. If the Delphi studies serve only to pinpoint the areas in which those administering the education, those receiving the education and those paying for the education differ, they will have been worth the cost and effort.

The methodology used to test the category two hypothesis tests was discussed in Chapter V. Briefly, the method involves calculating Ford's weights for the round three goals in those goal areas determined to have attained

consistency in preferences in section 6.5 of Chapter VI. For a comparison to be made of a goal area between two categories of judges, i.e., between two Delphi studies, each category of judges must have had consistent preferences in that goal area. Given a goal area with consistent preferences in two judge categories, we desire to test the null hypothesis that the judges in the two judge categories are from a common population of judges, i.e., the preferences of the judges in one judge category do not differ significantly from the preferences of the judges in the other category. Using the test statistic $\bar{d}$ and the randomization distribution, the probability of obtaining a worse value of $\bar{d}$ than was observed is calculated, i.e., the p value. Thus, an exact probability assessment is made concerning the likelihood of the two judge categories being from a common population of judges.

The application of a category two hypothesis test and a comparison of the results obtained from using several available methods of analysis will be covered in section 7.2. The results of all of the category two hypothesis tests conducted during this research will be examined in detail in the last section of the chapter.

7.1 Comparative Example

To illustrate the proposed methodology for testing category two hypothesis tests consider the following example.

Given are the preferences of two categories of judges for the consistent goal area 3.1.2, "Forming Relationships with Others," obtained during questionnaire three of each judge category. In this example the two judge categories are community leaders and students. The Ford weights for each goal judged in the two categories and the differences in the Ford weights for each goal in the two categories are located in Table 8. The value of the test statistic $\bar{d}$ is .498321 for the two categories of judges in this goal area. Generating the null distribution of $\bar{d}$, located in Figure 5, we obtain a p value of .187. Using a significance level of .10 for α, we fail to reject the null hypothesis of no differences in preference between categories of judges, i.e., no differences in preference between community leaders and students in this goal area. Upon closer examination of the differences used in calculating $\bar{d}$, we see that the community leaders perceived four of the five goals as being more important than did the students. On the fifth goal, the individual understands the judicial system, the two groups were almost identical in their perceptions. Clearly, there is nothing magical about the α value of .10 used in this example. Quite possibly, a more realistic value for this type of research would be an α of .20 or .25.

To provide a means of comparison, the Kruskal-Wallis test described in section 2.3 of Chapter II, will now be used to test the null hypothesis that the two samples are

Table 8. Forming Relationships with Others: Community Leaders Versus Students

Goal Statement	Community Leaders Round Three	Students Round Three	Differences Community Leaders-Students
Understands the judicial system	.067027	.075192	-.008165
Has concern for his fellow man	4.301439	2.319599	1.981840
Understands and respects the beliefs and feelings of others	*	2.901324	*
Understands and appreciates persons of different backgrounds as a result of personal experience with and exposure to many different kinds of persons	.458224	.243330	.214894
Knows how to live in an urban community so as to respect the physical, emotional and mental well-being of others	.487487	.231325	.256162
Knows how and wishes to create meaningful relationships with others	*	1.361890	*
Has the ability to evaluate the character of other individuals	.124719	.077845	.046874

*Indicates that the goal was not rated by the community leaders.

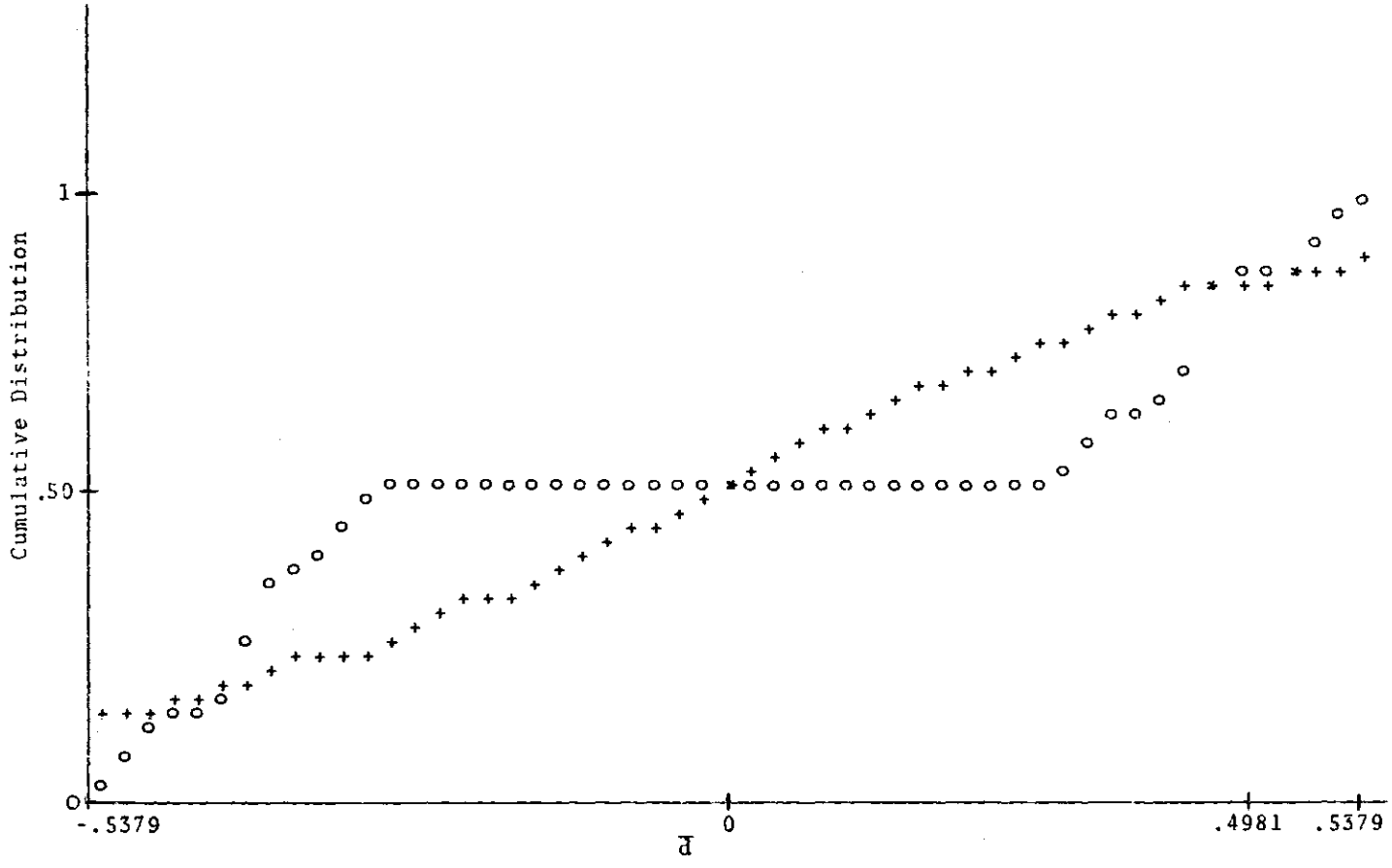


Figure 5. Forming Relationships with Others: Community Leaders Versus Students

from a common population. The Ford weights for the two categories of judges are ranked below from largest value to smallest value.

Community Leaders	Students
10	9
1	2
4	5
3	6
7	8
Rank Sum = 25 (R_1)	Rank Sum = 30 (R_2)

The value of the test statistic H is calculated as

$$H = \left(\frac{12}{n(n+1)}\right)\left(\frac{(R_1)^2}{n_1} + \frac{(R_2)^2}{n_2}\right) - 3(n+1) \qquad (1)$$

where n = total number of observations, n_1 = number of observations in the first sample, n_2 = number of observations in the second sample, R_1 = rank sum of sample one, and R_2 = rank sum of sample two. Hence,

$$H = \left(\frac{12}{10.11}\right)\left(\frac{(25)^2}{5} + \frac{(30)^2}{5}\right) - 3(11) = .273 \qquad (2)$$

and from the χ^2 tables with one degree of freedom we find that the probability of $H \geq .273$ is approximately .619, when the null hypothesis is true. Therefore, we fail to reject the null hypothesis at the $\alpha = .10$ level. Although we failed to reject the hypothesis, we lost much of the

resolution available in the observations by converting them to ranks.

If the t distribution is used to test the null hypothesis that the two samples are from a common population, the probability of obtaining a worse value for the test statistic than the value of .498321 observed is .262. Although the values from the t and the randomization distributions are relatively close for these two samples, in general this would not be true. Figure 5 shows that there are considerable differences in the two distributions and that quite different inferences would be drawn for most values of the test statistic $\bar{d}$.

7.2 Results of Category Two Hypothesis Tests

The complete results of the 27 category two hypothesis tests performed are listed in Table 10. Only those goal areas that contained consistent preferences in at least two judge categories were compared. Using a two-tailed test and a significance level of .10 leads to the conclusions outlined below. Table 9 contains a complete listing of goals by goal area in which significant differences were observed and the associated Ford weight differences used in calculating the test statistic used for each category two hypothesis test.

Community leaders and students and community leaders and educators do not differ significantly in their

Table 9. Goal Areas in Which the Category Two Null Hypothesis Was Rejected

Goal Area	Goal Number	Goal Statement	di's for C.L. - Stu.	di's for C.L. - Ed.	di's for Stu. - Ed.
1.0	83	Is able to listen, speak, read and write	.332666	2.635220	-9.758903*
1.0	10	Is able to communicate feelings, ideas, and information	12.443706	-1.065981	- .682596*
1.0	120	Is able to speak with ease in front of a large group	--	--	.290961*
1.0	95	Understands that basic skills in listening, speaking, reading, writing, and mathematics are necessary to be able to do most things of importance in modern society	--	--	-2.203406*
2.3	21	Possesses the attitudes and skills to pursue learning as a life-long process	-.602926	-.953153	- .348837*
2.3	105	Knows how to study effectively	--	--	- .433117*
2.3	67	Knows how and where to obtain additional training and education	.419947	-1.190439	-1.489152*
2.3	30	Possesses the ability and desire to use the learning resources of the community	.250873	.232841	- .017515*
3.1.1	74	Recognizes that every man has the right to participate freely in society as long as the rights of others are not violated	3.954804	.077890*	-3.573636*

*Indicates that these d_i's were significantly different within each goal area for the categories tested.

Table 9. (Continued)

Goal Area	Goal Number	Goal Statement	di's for C.L. - Stu.	di's for C.L. - Ed.	di's for Stu. - Ed.
3.1.1	104	Understands how freedom is related to self-discipline	--	*	-2.121823*
3.1.1	9	Understands and accepts the relationship of rights to responsibilities	4.693976	-2.516380*	-6.680336*
3.1.1	82	Understands freedom as the right to make choices within the framework of concern for the general welfare	-1.455022	-2.487929*	- .907055*
3.1.1	72	Is committed to the values expressed in the Bill of Rights	2.263706	.918060*	-1.169961*
3.1.1	71	Has the desire to preserve the rights and property of others	1.329068	-6.496697*	-7.013791*
3.1.1	86	Supports the free and voluntary exercise of religious choice	-1.313974	-1.036399*	.161723*
3.1.1	6	Understands and accepts the necessity and desirability of avoiding discrimination in employment practices	-2.471570	-1.281155*	1.004324*
3.1.1	19	Is willing to live in a racially Integrated society	- .502755	- .354354*	.103110*
3.1.1	23	Is able to understand and tolerate dissent	-1.987011	.388012*	2.118473*
3.1.1	114	Has a sense of personal freedom, particularly freedom of expression	.238069	.472068*	.187359*
3.1.1	56	Understands and is committed to the processes and purposes of law	- .586602	- .810352*	- .222319*
3.1.1	11	Understands the judicial system	- .008165	.060294*	.058041*

Table 9. (Continued)

Goal Area	Goal Number	Goal Statement	di's for C.L. - Stu.	di's for C.L. - Ed.	di's for Stu. - Ed.
4.2	47	Has the knowledge and skills for managing personal finances	-1.095633	-1.784544*	- .635751
4.2	40	Has the knowledge and skills to be successful in meeting his needs as a consumer of goods and services	.014735	.002683*	- .012384
4.3	60	Is able to secure information from a wide variety of sources, to analyze, to synthesize, to draw conclusions and to make decisions	.221834*	- .677736	- .829531
4.3	116	Is able to ask meaningful questions and to evaluate the answers he gets	*	--	- .035980
4.3	111	Recognizes the difference between believing something because of who says it and believing it because of its true value	1.122923*	1.323102	.169169
4.3	91	Has the ability to objectively evaluate issues, concepts, problems, etc.	.290916*	.460441	.133252
4.3	34	Is able to make responsible decisions regarding the use of time	.419547*	.393232	- .023167
5.1	24	Possess knowledge, understanding and appreciation of his heritage	- .437265	- .374024	.041888*
5.1	117	Understands how historical events and patterns are related to present events and patterns	--	--	.040764*

Table 9. (Continued)

Goal Area	Goal Number	Goal Statement	di's for C.L. - Stu.	di's for C.L. - Ed.	di's for Stu. - Ed.
5.1	103	Understands the part played in our society by mass media--newspapers, magazines, radio, TV, etc.	--	--	.016422*
5.1	69	Understands the functions of public education in our society and how it is administered	- .793167	.017310	.699109*
5.1	110	Has a basic knowledge of human psychology (how the human mind and emotions work)	--	--	.090704*
5.1	35	Recognizes the influence of the family and religious and community organizations in shaping values in a changing society	- .040481	.046191	.068549*
5.1	33	Knows and understands the concepts of taxation	- .030562	.040732	.059454*
5.1	36	Understands the structure and functions of local, state and national governments	.047249	.083642	.030481*
5.1	42	Understands and appreciates the contributions of social, religious and national groups to our culture	.016888	.031131	.011834*
5.1	77	Has knowledge of the principle economic, social and political systems of the world	- .004560	.025225	.025474*
5.1	41	Has a knowledge and understanding of international relations	- .006718	.006847	.011656*

Table 9. (Concluded)

Goal Area	Goal Number	Goal Statement	di's for C.L. - Stu.	di's for C.L. - Ed.	di's for Stu. - Ed.
5.1	66	Possesses knowledge and understanding of production, distribution, and consumption of agricultrual and industrial products	.001838	.003837	.001376*
5.1	25	Has a knowledge and understanding of current political issues	.041822	.060042	.014624*

*Indicates that these d_i's were significantly different within each goal area for the categories tested.

Table 10. Results of Category Two Hypothesis Tests

$Pr(|\bar{d}| \geq \text{observed value}) = p$. (The figures represent values of p.)

Goal Areas	C.L. vs Stu.	C.L. vs Ed.	Stu. vs Ed.
Communication Skills	.846	.848	.084
Self-Understanding	*	.105	*
Career Development	.357	.344	.771
Preparation for Life-Long Learning	.956	.348	.063
Preparation for Leisure	*	*	.224
Commitment to the Principles of Democracy	.647	.049	.090
Forming Relationships With Others	.187	.956	.344
Participating Actively as a Citizen	*	*	*
Preparation for Family Life	.848	.794	.501
Life Skills	**	**	**
Preparation for Managing Health and Environment	*	*	*
Preparation for Managing Personal Finances	.457	.043	.250
Problem Solving	.100	.453	.773
Social Sciences	*	*	.001
Mathematics	**	**	**
Science and Technology	*	*	*

*No test was performed since the goal area did not converge to consistency in both delphi studies.

**No test was performed since the goal area only contained one goal.

perceptions on the importance of Communication Skills, goal area 1.0, however, students and educators do differ significantly in their perceptions on the importance of Communication Skills. It must be pointed out that the community leaders did not rate two of the four goals contained in the Communications Skills area, and students and educators rated all four goals. The significant difference between students and educators is due to their preferences on the two goals not rated by community leaders, i.e., goal numbers 95 and 120. Examination of the Ford weights assigned to each goal in the student Delphi study and in the educator Delphi study shows that the educators perceived three of the four goals as being more important than the students did, i.e., goal numbers 83, 10 and 95 were perceived as more important by the educators. The greatest difference in opinion occurred on goal number 83, the individual is able to listen, speak, read and write. The educators rated that goal as most important of the 121 goals, but the students rated it as tenth in overall importance.

Community leaders versus educators was the only hypothesis tested in goal area 2.1, Self-Understanding, since the students did not have consistent preferences in this area during their Delphi study. The p value obtained from the randomization distribution was .105. Whether the null hypothesis should be rejected is questionable in this

goal area. It is obvious that the two groups differ in their preferences. Analysis of the Ford weights for the seven goals in this area reveals that the two groups agreed on one goal, goal number 85, and that the educators perceived five of the remaining six goals as being more important than the community leaders did.

In goal area 2.2, Career Development, the null hypothesis was not rejected for any of the tests performed. There was no significant differences in preference between any two categories of judges.

Community leaders and students are in close agreement in the perceived importance of the three goals in goal area 2.3, Preparation for Life-Long Learning. Community leaders and educators can be considered as being in agreement; however, they do not appear to agree as much as community leaders and students do. In testing students versus educators we must reject the null hypothesis of no preference difference at the $\alpha = .10$ level. Looking at the differences in Ford weights for the two groups, we find that the educators perceived all of the goals to be of more importance than the students did. Although there was one additional goal in this area rated by the students and educators, goal number 105, it did not appear to be the deciding influence in rejecting the null hypothesis. It seems that the preferences of the community leaders were between those of the educators and the students and thus were

not significantly different from either one.

Since the community leaders did not have consistent preferences in their Delphi study, the only test made in goal area 2.4, Preparation for Leisure, was between students and educators. The randomization distribution yielded a p value of .224. The two groups are in partial agreement in this goal area. The differences in Ford weights indicate that the two groups are in agreement on one goal, goal number 27, and that the students rated all but one, goal number 68, of the remaining goals as being more important than the educators did.

Community leaders and students appear to be in agreement in goal area 3.1, Preparation for Citizenship. Tests were conducted in subareas 3.1.1 and 3.1.2, Commitment to the Principles of Democracy and Forming Relationships with Others, respectively. Though we fail to reject the null hypothesis in both subareas, we have more confidence in our decision in the first subarea. In the tests between community leaders and educators, the null hypothesis is rejected in the subarea of Commitment to the Principles of Democracy. However, in the area of Forming Relationships with Others we confidently fail to reject the null hypothesis. When students and educators are compared, the null hypothesis is rejected in the first subarea, 3.1.1, but not in the second subarea, 3.1.2. Examination of the Ford weights for the goals rated by students and educators in the

first subarea shows that the educators perceived four of the 12 goals as being significantly more important than the students did. These four goals are numbers 74, 104, 9 and 71.

In the area of Preparation for Family Life, we fail to reject the null hypothesis in each of the three tests. Since all three values in the table are greater than .50, the null hypothesis is highly tenable.

In goal area 4.2, Preparation for Managing Personal Finances, the null hypothesis of no preference difference is rejected when comparing the preferences of community leaders and educators. However, when comparing the preferences of community leaders and students and students and educators we fail to reject the null hypothesis. The reason for rejecting the null hypothesis when comparing community leaders and educators seems to be due to goal number 47, the Individual has the Knowledge and Skills for Managing Personal Finances. Community leaders rated this goal significantly higher than did the educators.

When comparing community leaders and educators and students and educators in the area of Problem Solving, goal area 4.3, we fail to reject the null hypothesis of no preference difference. However, when comparing community leaders and students, the p value was found to be .100 which is also the α level. Since a high level of consistency in preference was not achieved in the student Delphi study in

this goal area, a p value of .250 on the final comparison was obtained, the value of the information is marginal before the hypothesis test is made. Based on that fact, this goal area certainly requires more investigation before a definitive statement can be made regarding the null hypothesis.

In the area of Social Sciences, goal area 5.1, the only test conducted was between students and educators, since the community leaders were not consistent in their preference in this area in their Delphi study. The null hypothesis is rejected since the p value is .001. Examination of the differences in the Ford weights for each goal in these two categories of judges shows that the students rated all 13 goals in this area higher than did the educators.

There were five goal areas in which no category two hypothesis tests were performed. In two areas, Knows How to Secure and Use Community Services and Mathematics, no test was made, since each area only contained one goal. No test was made in the following three areas since they did not possess consistent preferences in at least two of the Delphi studies; (1) Participating Actively as a Citizen, (2) Preparation for Managing Health and Environment: Physical and Mental Health, and (3) Science and Technology.

In summary, category two hypothesis tests were made in those goal areas in which the Delphi technique was successful in obtaining consistent preferences. That is, those areas in which the probabilistic assessment did not

allow the rejection of the null hypothesis. Only when a category of judges' preferences are consistent can we assert that the Delphi technique has been successful, and only when the Delphi technique has been successful is the information obtained of sufficient value for drawing future statistical inferences.

The results of the category two hypothesis tests indicate that community leaders and students are in closest agreement with respect to their preferences for educational goals. They only disagreed significantly in one area, goal area 4.3; and the p value for this hypothesis test was .100.

The community leaders and educators differed significantly in only two goal areas, 3.1.1 and 4.2. Overall, the educators felt that the goals contained in goal area 3.1.1, Commitment to the Principles of Democracy, were of higher importance than did the community leaders. In goal area 4.2, Preparation for Managing Personal Finances, the two felt the same about goal number 40, but the educators felt that the remaining goal, number 47, was considerably more important than did the community leaders.

Educators and students disagreed significantly in four goal areas. These areas were Communication Skills, Preparation for Life-Long Learning, Commitment to the Principles of Democracy and Social Sciences. In the first three areas listed the educators felt that 14 of the 20 goals contained in these areas were of higher importance

than did the students. However, in the area of Social Sciences just the opposite was observed, i.e., the students felt that all 13 goals contained in this area were of higher importance than did the educators.

The category two hypothesis tests have been successful in identifying those areas in which significant differences in preference exist between categories of judges. The Ford weights have allowed for an indication of the direction of disagreement and also for a resolution of which goal or goals within an area is responsible for the observed difference.

CHAPTER VIII

CONCLUSIONS

The methodology proposed in this research for analyzing ordinal data has been shown to be an efficient method of analysis. The criterion of consistency in preference is a viable alternative to the various criteria dealing with opinion convergence in ascertaining the success of a Delphi study, i.e., in ascertaining the value of the information obtained from a Delphi study. The consistency criterion is applicable to any Delphi study in which value judgments are concerned. The Ford procedure provides the resolution necessary to determine the relative magnitude of importance associated with each item being rated. The randomization distribution is the exact distribution for the observed sample; therefore, no assumptions regarding the underlying distribution of the judges' preferences or the group's preferences are required. The probabilistic assessments obtained from it are easily interpreted in the context of the Delphi technique, and they are as powerful as any of the nonparametric or parametric tests applicable to this type data.

The category one hypothesis tests identified those goal areas in which consistent preferences were obtained,

i.e., those goal areas in which valuable information is contained. These results provide the experimenter with a convenient method of evaluating the structure of the instrument used in the Delphi study and of evaluating the effectiveness of the feedback information utilized during the Delphi study. The category two hypothesis tests were successful in identifying specific goal areas in which each category of judges differed significantly in their preferences for the educational goals. More importantly, the Ford weights used in these tests indicate the direction of the preferences that exist within particular goal areas.

The category one hypothesis tests indicated that the Delphi technique was successful in obtaining a category of judges' consistent preferences in each Delphi study in the following goal areas: Communication Skills, Career Development, Preparation for Life-Long Learning, Commitment to the Principles of Democracy, Forming Relationships with Others, Preparation for Family Life, Preparation for Managing Personal Finances and Problem Solving. The implications of the consistency in preference obtained in these goal areas are: (a) all three categories of judges were able to formulate stable preferences in these eight areas; (b) the respondents have definite perceptions on the importance of the goals contained in these areas; and (c) these areas contain the only information of value for performing subsequent hypothesis tests between the categories of judges.

The areas of Preparation for Managing Health and Environment, Science and Technology, Life Skills and Mathematics do not contain information that can be considered valuable enough to use in making any statistical inferences. The areas either did not contain consistent preferences or they did not contain a sufficient number of goals to allow for any measurable resolution of the data. The specific goals within these areas should be analyzed to see what effect the absence of data indicating their perceived importance may have on future and present priorities in the education system of Atlanta. If their input is deemed essential to the development of future educational objectives, a followup study concentrating on these areas would be in order.

In the category two hypothesis tests it was observed that community leaders and students differed significantly in only one area, Problem Solving. Community leaders and educators differed significantly in two goal areas, Commitment to the Principles of Democracy and Preparation for Managing Personal Finances. Most importantly, educators and students differed significantly in the following four areas: (a) Communication Skills, (b) Preparation for Life-Long Learning, (c) Commitment to the Principles of Democracy and (d) Social Sciences. The differences between educators and students require immediate attention.

Obviously, if those implementing the educational process perceive educational goals differently than those

receiving the instruction, the motivation and attentiveness of both will ultimately be affected.

The fact that students perceive goal number 90, the Individual Has a Marketable Skill When he Leaves High School, even though he may plan to go on for further education, as being only moderately important indicates that they do not feel a need for more career education courses. Community leaders placed a higher value of perceived importance on this goal and in general on the entire area of Career Development than did students or educators. The implications are that the respondents are more in favor of a general education program at the secondary level. Perhaps this is due in large part to the various postsecondary vocational opportunities available in the Atlanta area.

APPENDICES

APPENDIX A

CATEGORIES OF THE TAXONOMY OF EDUCATIONAL OBJECTIVES

Cognitive Domain

KNOWLEDGE

1.00 KNOWLEDGE

1.10 KNOWLEDGE OF SPECIFICS

1.11 KNOWLEDGE OF TERMINOLOGY

1.12 KNOWLEDGE OF SPECIFIC FACTS

1.20 KNOWLEDGE OF WAYS AND MEANS OF DEALING WITH SPECIFICS

1.21 KNOWLEDGE OF CONVENTIONS

1.22 KNOWLEDGE OF TRENDS AND SEQUENCES

1.23 KNOWLEDGE OF CLASSIFICATIONS AND CATEGORIES

1.24 KNOWLEDGE OF CRITERIA

1.25 KNOWLEDGE OF METHODOLOGY

1.30 KNOWLEDGE OF THE UNIVERSALS AND ABSTRACTIONS IN A FIELD

1.31 KNOWLEDGE OF PRINCIPLES AND GENERALIZATIONS

1.32 KNOWLEDGE OF THEORIES AND STRUCTURES

INTELLECTUAL ABILITIES AND SKILLS

2.00 COMPREHENSION

2.10 TRANSLATION

2.20 INTERPRETATION

2.30 EXTRAPOLATION

3.00 APPLICATION

4.00 ANALYSIS

4.10 ANALYSIS OF ELEMENTS

4.20 ANALYSES OF RELATIONSHIPS

4.30 ANALYSIS OF ORGANIZATIONAL PRINCIPLES

5.00 SYNTHESIS

5.10 PRODUCTION OF A UNIQUE COMMUNICATION

5.20 PRODUCTION OF A PLAN OR PROPOSED SET OF OPERATIONS

5.30 DERIVATION OF A SET OF ABSTRACT RELATIONS

6.00 EVALUATION

6.10 JUDGMENTS IN TERMS OF INTERNAL EVIDENCE

6.20 JUDGMENTS IN TERMS OF EXTERNAL CRITERIA

APPENDIX B

CATEGORIES OF THE EFFECTIVE DOMAIN OF THE TAXONOMY OF EDUCATIONAL OBJECTIVES

1.0 RECEIVING (ATTENDING)

- 1.1 AWARENESS
- 1.2 WILLINGNESS TO RECEIVE
- 1.3 CONTROLLED OR SELECTED ATTENTION

2.0 RESPONDING

- 2.1 ACQUIESCENCE IN RESPONDING
- 2.2 WILLINGNESS TO RESPOND
- 2.3 SATISFACTION IN RESPONSE

3.0 VALUING

- 3.1 ACCEPTANCE OF A VALUE
- 3.2 PREFERENCE FOR A VALUE
- 3.3 COMMITMENT

4.0 ORGANIZATION

- 4.1 CONCEPTUALIZATION OF A VALUE
- 4.2 ORGANIZATION OF A VALUE SYSTEM

5.0 CHARACTERIZATION BY A VALUE OR VALUE COMPLEX

- 5.1 GENERALIZED SET
- 5.2 CHARACTERIZATION

APPENDIX C

A CONDENSED VERSION OF THE PSYCHOMOTOR DOMAIN OF THE TAXONOMY OF EDUCATIONAL OBJECTIVES

1.0 PERCEPTION

- 1.1 SENSORY STIMULATION
  - 1.11 AUDITORY
  - 1.12 VISUAL
  - 1.13 TACTILE
  - 1.14 TASTE
  - 1.15 SMELL
  - 1.16 KINESTHETIC
- 1.2 CUE SELECTION
- 1.3 TRANSLATION

2.0 SET

- 2.1 MENTAL SET
- 2.2 PHYSICAL SET
- 2.3 EMOTIONAL SET

3.0 GUIDES RESPONSE

- 3.1 IMITATION
- 3.2 TRIAL AND ERROR

4.0 MECHANISM

5.0 COMPLEX OVERT RESPONSE

- 5.1 RESOLUTION OF UNCERTAINTY
- 5.2 AUTOMATIC PERFORMANCE

APPENDIX D

A SOURCE LISTING AND AN EXPLANATION OF FORD'S PROCEDURE

Ford rank, a binary comparison ranking program, accepts ranks of N objects assigned by K judges, where not all K judges rank all objects and where each judge may use a different number of ranks. Objects universally rated high or low are scanned and omitted from the calculation. A win-loss paired comparison matrix is calculated and a set of normalized weights is produced for each object corresponding to its ranked position relative to all others being considered. The number of objects ranked can not be greater than 130 nor can the number of ranks exceed 130.

There will be five different data cards for each matrix or quality judged: (1) Label Card, (2) Parameter Card, (3) Judge Card, (4) Rank Card A, and (5) Rank Card B--optional.

Columns 73-80 of all types of cards are not used by the program but are used for ordering the cards; they should be coded as follows:

Columns

73-74 Matrix Number

01-XX All cards: matrix number; lead zero must be punched. Since only one matrix is run at a time, it

does not matter whether this number begins with 01 or is sequential.

75 Card Type

1 Label Card

2 Parameter Card

3 Judge Card, Rank Card A, Rank Card B: Judge number; lead zero must be punched. (These should start with 01 and be sequential.)

76-77 Judge Number

00 Label Card, Parameter Card, Judge Card

01-XX Rank Card A, Rank Card B: rank number; lead zero must be punched. (These should start with 01 and be sequential.)

80 Type of Rank Card

0 Label card, Parameter card, Judge card

1 Rank card A

2 Rank card B

LABEL CARD (one for each matrix)

1 Must be 1 (one)

2-72 Any information; this will be printed out by the machine.

73-80 See preceding information.

PARAMETER CARD (one for each matrix)

1-6 Total number of people being compared by all judges. This cannot exceed 130. Omit lead zeros; last digit must be in column 6.

7-12 Number of judges. Omit lead zeros; last digit must be in column 12.

13-18 Criterion for convergence; omit zeros preceding the decimal; last digit must be in column 18. When the weights generated by the program for each individual change less than the amount punched here from one iteration to the next, the problem is said to have converged and iterations stop. (For this research a convergence criterion of .005 was used.)

19-24 The maximum number of iterations. Omit lead zeros; last digit must be in column 24. (For this research 200 was used.)

25-72 Blank

73-80 See preceding information.

JUDGE CARD (One for each judge for each matrix)

1-6 Number of ranks used by this judge. Omit lead zeros; last digit must be in column 6.

7-72 Blank

73-80 See preceding information.

RANK CARD A (one for each rank used by each judge for each matrix)

1-3 Number of people placed in this rank by this judge. Omit lead zeros; last digit must be in column 3.

4-6 I.D. number of person placed in this rank. Do not omit lead zero.

7-9 It will be noted that the last digit of the I.D.

10-12 number always falls in a column number which is
..... divisible by 3. (If more than 23 people are placed in
67-69 this rank by this judge, a Rank card B will be needed.)
70-72 If less than 23 people are placed in this rank by this judge, blanks will occur for the remainder of the columns through column 72.

I.D. numbers of persons judged do not need to be within the range of 001-130 or in ascending order. The program will assign an "assigned I.D. number" to each "original I.D. number." A "map" will be printed out containing both numbers. The "assigned I.D. number" will be assigned from 001-130 in the order in which the "original I.D. numbers" are encountered in the data. The final weights will be printed out in order of the "assigned I.D. number" but the "original I.D. number" will also be printed.

73-80 See preceding information.

RANK CARD B (occurs whenever a judge puts more than 23 people in one rank; there may be additional Rank card B's if needed. If 23 or fewer people are put in this rank, Rank card B is omitted.)

1-3 I.D. number of person placed in this rank. (See Rank card A for additional comments.)

4-6

7-9

.....

67-69

70-72

73-80 See preceding information.

SETUP

1. Program

2. Label Card

3. Parameter Card

4. Judge Card of first judge

5. Rank cards of the first judge in rank order; i.e. the rank card(s) for rank one followed by the successive rank cards for the successive ranks. Rank card A precedes Rank card B if Rank card B occurs; if there are two or more Rank card B's their order does not matter.

-- Repeat 4 and 5 for remainder of judges in order by ascending judge number.

-- 2-5 may be ordered by regarding the numbers in column 73-80 as one number and arranging the cards in order in ascending sequence. The output consists of:

A list of the judges and the number of comparisons made.

An I.D. number map containing assigned I.D. numbers and original I.D. numbers.

A win-loss matrix ordered by assigned I.D. number.

A list of people assigned I.D. number rated universally low or high.

The iterations.

A list of final weights in order of assigned I.D. numbers, and listing also the original I.D. numbers.

```
      COMMON LIST(130)
      COMMON NG,NGM1,NTGJ,NTG(130),EPSLON,NR(130),NC(130),N,NCOUNT,JUPPE
     1R,MC(130,130),M1,M2,INFW,JNEW,IP1,W(130),A(130,130),X(130),MAN(130
     2),JO,KO,DENOM,XA,XB,Y(130)
      INTEGER OUTPUT
      LEN=0
      INPUT=5
      OUTPUT=6
      READ(5,1)JJ,JUPPER,N,EPSLON
    1 FORMAT(3I3,1F6.3)
      NCOUNT=0
      DO 8 J1=1,JJ
      NCNT1=NCOUNT
      READ(5,102)NG
  102 FORMAT(1I6)
  112 DO 5 NG1=1,NG
      READ(5,103)NTG1,(MC(NG1,K5),K5=1,NTG1)
  103 FORMAT(4I3)
      DO 10 I=1,NTG1
      IF(LEN.EQ.0) GO TO 25
      DO 20 J=1,LEN
      IF(MC(NG1,I).EQ.LIST(J)) GO TO 30
   20 CONTINUE
   25 LEN=LEN+1
      LIST(LEN)=MC(NG1,I)
      J=LEN
   30 MC(NG1,I)=J
   10 CONTINUE
    5 NTG(NG1)=NTG1
      NGM1=NG-1
      DO 7 I=1,NGM1
      IP1=I+1
      NTGI=NTG(I)
      DO 7 J=IP1,NG
      NTGJ=NTG(J)
      DO 7 IC=1,NTGI
      DO 7 JC=1,NTGJ
      NN1=MC(I,IC)
      NN2=MC(J,JC)
      IF(NN1-NN2)6,7,6
    6 A(NN1,NN2)=A(NN1,NN2)+1.
      NCOUNT=NCOUNT+1
      NR(NN1)=1
      NC(NN2)=1
    7 CONTINUE
      NCNT1=NCOUNT-NCNT1
    8 WRITE(OUTPUT,9)J1,NCNT1
    9 FORMAT(6H0JUDGE,1I5,10(1H ),1I7,12H COMPARISONS )
      WRITE(OUTPUT,500)
  500 FORMAT(1H0,'ID NUMBER MAP')
      WRITE(OUTPUT,501)
  501 FORMAT(15(1H ),13HASSIGNED ID #,15(1H ),13HORIGINAL ID #/)
      WRITE(OUTPUT,502)(I,LIST(I),I=1,N)
  502 FORMAT(T20,1I4,T49,1I4)
```

```
      NR(I)=0
   31 NC(I)=0
      WRITE(OUTPUT,400)NCOUNT
  400 FORMAT(1H0,'NCOUNT =',I10)
      CALL CORE2
      CALL CORE3
      END
      SUBROUTINE CORE2
      COMMON LIST(130)
      COMMON NG,NGM1,NTGJ,NTG(130),EPSLON,NR(130),NC(130),N,NCOUNT,JUPPE
     1R,MC(130,130),M1,M2,INEW,JNEW,IP1,W(130),A(130,130),X(130),MAN(130
     2),JO,KO,DENOM,XA,XG,Y(130)
      INTEGER OUTPUT
      OUTPUT=6
      INPUT=5
      ITER=0
      DO 34 I=1,N
   34 MAN(I)=I
   33 DO 3 I=1,N
      IF(NR(I)*NC(I))3,4,3
    3 CONTINUE
      WRITE(OUTPUT,301)
  301 FORMAT(1H0,'WIN-LOSS MATRIX ORDERED BY ASSIGNED ID NUMBER')
      DO 70 I=1,N
   70 WRITE(OUTPUT,36)(A(I,J),J=1,N)
   36 FORMAT(26F5.0)
      GO TO 100
    4 INEW=0
      WRITE(OUTPUT,300)
  300 FORMAT(1H0,'WIN-LOSS MATRIX ORDERED BY ASSIGNED ID NUMBER')
      DO 71 I=1,N
   71 WRITE(OUTPUT,37)(A(I,J),J=1,N)
   37 FORMAT(26F5.0)
      ITER=ITER+1
      DO 10 I=1,N
      IF(NR(I)*NC(I))9,18,9
    9 INEW=INEW+1
      JNEW=1
      MAN(INEW)=MAN(I)
      DO 1000 J=1,N
      IF(NR(J)*NC(J))11,1000,11
   11 A(INEW,JNEW)=A(I,J)
      JNEW=JNEW+1
 1000 CONTINUE
      GO TO 10
   18 IF(NR(I))8,5,8
    8 WRITE(OUTPUT,12)MAN(I),ITER
   12 FORMAT(1H0,22HSUBJECT ASSIGNED ID #,I3,35H IS UNIVERSALLY RATED HI
     1GH,DELETED,I3, 3H RY )
      GO TO 10
    5 WRITE(OUTPUT,7)MAN(I),ITER
    7 FORMAT(1H0,22HSUBJECT ASSIGNED ID #,I3,35H IS UNIVERSALLY RATED LO
     1W ,DELETED,I3, 3H RY )
   10 CONTINUE
      N=INEW
      DO 31 I=1,N
```

```
      DO 30 I=1,N
      DO 30 J=1,N
      IF(A(I,J))32,30,32
   32 NR(I)=1
      NC(J)=1
   30 CONTINUE
      GO TO 33
  100 DO 16 I=1,N
      W(I)=0.
      DO 13 J=1,N
   13 W(I)=W(I)+A(I,J)
      Z=0.
      DO 15 J=1,N
   15 Z=Z+A(J,I)
   16 X(I)=W(I)/(W(I)+Z)
      DO 200 I=1,N
      DO 200 J=1,N
  200 A(I,J)=A(I,J)+.00001
      DO 14 I=1,N
      IP1=I+1
      DO 14 J=IP1,N
      A(I,J)=A(I,J)+A(J,I)
   14 A(J,I)=A(I,J)
      RETURN
      END
      SUBROUTINE CORE3
      COMMON LIST(130)
      COMMON NG,NGM1,NTGJ,NTG(130),EPSLON,NR(130),NC(130),N,NCOUNT,JUPPE
     1R,MC(130,130),M1,M2,INEW,JNEW,IP1,W(130),A(130,130),X(130),MAN(130
     2),JO,KO,DENOM,XA,XQ,Y(130)
      INTEGER OUTPUT
      OUTPUT=6
      INPUT=5
      JO=0
    8 KO=0
      DO 3 I=1,N
      DENOM=0.
      DO 4 J=1,N
    4 DENOM=DENOM+A(I,J)/(X(I)+X(J))
    3 Y(I)=A(I)/DENOM
      JO=JO+1
    5 DO 9 I=1,N
      IF(ABS(Y(I)-X(I))/X(I)-EPSLON)9,9,10
   10 KO=KO+1
    9 X(I)=Y(I)
      WRITE(OUTPUT,12)JO,KO,(Y(I),I=1,N)
   12 FORMAT(1H0,I5,I10/(6F18.6))
      IF(KO)17,17,13
   13 IF(JO-JUPPER)8,14,14
   14 WRITE(OUTPUT,15)JUPPER
   15 FORMAT(1H0,I5,46H ITERATIONS, NO CONVERGENCE. DATA SET DELETED. )
   17 WRITE(OUTPUT,18)
   18 FORMAT(14H1FINAL WEIGHTS//15(1H ),
     113HASSIGNED ID #, 3(1H ),13HORIGINAL ID #,12(1H ),6HWEIGHT//)
      DO 99 KK=1,N
      II=MAN(KK)
   99 WRITE(OUTPUT,19)II,LIST(II),Y(KK)
   19 FORMAT(1H0,2I19,F24.6)
      END
```

BIBLIOGRAPHY

1. Andrews, F. M. and Pelz, D. C., "Ford's Procedure for Combining Multiple Sets of Partially Ordered Data," Educational and Psychological Measurement, Vol. 33, No. 1 (1973), pp. 43-52.

2. Asch, S. E., "Effects of Group Pressure Upon the Modification and Distortion of Judgments." In Readings in Social Psychology, 3rd ed, Eleanor E. Maccoby, et al., eds. London: Holt, Rinehardt and Winston, 1958.

3. Atlanta Assessment Project, "Goals for Education in Atlanta, 1985--as seen by Community Leaders, Educators, and Students," April, 1973.

4. Betz, R. L., Engle, K. B. and Mallinson, G. G., "Perceptions of Non-College Bound, Vocationally Oriented High School Graduates," Personnel and Guidance Journal, Vol. 47, No. 1 (1969), pp. 988-994.

5. Bloom, B. S. and others, "Taxonomy of Educational Objectives, Handbook I: Cognitive Domain," David McKay Company, Inc., New York, New York, 1956, 207 pages.

6. Bloom, J. A., "Alternatives in Education--An Old Promise Renewed," Independent School Bulletin, Vol. 32, May 1973, pp. 27-29.

7. Bradley, James V., Distribution-Free Statistical Tests, Prentice-Hall, Englewood Cliffs, New Jersey, (1968).

8. Briggs, T. H., "The Basic Duties of a Secondary School," The National Association of Secondary School Principals Bulletin, Vol. 55, Pt. 2 (1971), pp. 84-88.

9. Brown, B. and Helmer, O., "Improving the Reliability of Estimates Obtained from a Consensus of Experts," The RAND Corporation, P-2986, September, 1964.

10. Cyphert, F. R. and Gant, W. L., "The Delphi Technique: A Case Study, "Phi Delta Kappa, Vol. 52 (1971), pp. 272-273.

11. Dalkey, N. C., "Delphi," The RAND Corporation, P-3704, October, 1967.

12. Ebel, R. L., "Evaluation and Educational Objectives," Journal of Educational Measurement, Vol. 10, No. 4 (1973), pp. 273-279.

13. Editorial Staff, "Schools Do A Poor Job in Teaching Noncollege Pupils," Nation's Schools, Vol. 87, No. 3 (1971), p. 41.

14. Editorial Staff, "Off-Campus Experiences Now an Approved Course of Study for High Schools," Nation's Schools, Vol. 90, No. 4 (1972), pp. 58-60.

15. Evans, R. N., "The Secondary School and Occupational Preparation," The National Association of Secondary School Principals Bulletin, Vol. 53, Pt. 1 (1969), pp. 23-40.

16. Fenner, J. L., "Reconnection for Relevance: A Proposed New High School Curriculum," The Record, Vol. 71, February (1970), pp. 423-438.

17. Finch, A. V., "Behavioral Analysis: Cognitive, Affective, and Psychomotor Domains," National Business Education Yearbook, No. 9 (1971), pp. 277-285.

18. Fisher, R., Design of Experiments, Hafner Publishing Company, Inc., 1960, 245 pages.

19. Ford, L. R., Jr., "Solution of a Ranking Problem from Binary Comparisons," The American Mathematical Monthly, Vol. 64, Pt. II, No. 8 (1957), pp. 28-33.

20. Foshay, A. W., "How Fare the Disciplines," The Phi Delta Kappan, Vol. 51, March 1970, p. 352.

21. Friedman, Milton, "A Comparison of Alternative Tests of Significance for the Problem of m Rankings," Annals of Mathematical Statistics, Vol. 11, 1940, pp. 86-92.

22. Galton, F., "Family Likeness in Stature," Proceedings of the Royal Society of London, Vol. 45, (1888), pp. 42-63.

23. Goodlad, J. I., "Curriculum: State of the Field," Review of Educational Research, Vol. 39, No. 3 (1969), pp. 367-375.

24. Grobman, H., "Curriculum Development and Evaluation," The Journal of Educational Research, Vol. 64, No. 10 (1971), pp. 436-442.

25. Hogan, R. F., "You'll Like It--It's Cannelloni!", Phi Delta Kappan, April, 1971, pp. 468-470.

26. Judd, R. C., "Delphi Method: Computerized 'Oracle' Accelerated Consensus Formation," College and University Business, September 1970, pp. 30-34.

27. Kelly, H. H. and Thibaut, J. W., "Experimental Studies of Group Problem Solving and Process," In Handbook of Social Psychology, Vol. II, Gardner Lindzey, ed. Reading, Massachusetts: Addison-Wesley Publishing Co., 1954.

28. Kendall, M. G., "A New Measure of Rank Correlation," Biometrika, Vol. 30, (June 38-January 39), pp. 81-93.

29. Kendall, M. G. and Smith, B. B., "The Problem of m Rankings," Annals of Mathematical Statistics, Vol. 10, 1939, pp. 275-287.

30. Kendall, M. G., Rank Correlation Methods, Charles Griffin & Company, Limited, 1970, London, England, 202 pages.

31. Krathwohl, D. R., Bloom, B. S., and Masia, B. B., "Taxonomy of Educational Objectives, Handbook II: Affective Domain," David McKay Company, Inc., 1964, New York, New York, 196 pages.

32. Kruskal, W. H. and Stuart, A., "Use of Ranks in One-Criterion Variance-Analysis," Journal of the American Statistical Association, Vol. 47, (1952), pp. 583-621.

33. Letson, J. W., "The Atlanta Assessment Project: A Plan for Operation," August, 1972.

34. Lieberson, Stanley, "Limitations in the Application of Nonparametric Coefficients of Correlation," American Sociological Review, October 1964, Vol. 29, Number 5, pp. 744-746.

35. Livingston, J. H., "Educational Goals and Graduate Opinions," Journal of Secondary Education, Vol. 44, No. 3 (1969), pp. 116-120.

36. Marland, S. P., Jr., "Career Education Now," The Vocational Guidance Quarterly, Vol. 20 (1972), pp. 188-192.

37. Miller, W. C., "Career Education and the Curriculum Leader," The Education Digest, February 1974, No. 1, pp. 48-50.

38. National Commission on the Reform of Secondary Education, "Rationale for a New Examination of Secondary Education," The Education Digest, February 1974, No. 1, pp. 2-5.

39. Ohme, H., "Steps Toward Relevance: An Interest-Centered Curriculum," Journal of Secondary Education, Vol. 45, No. 7 (1970), pp. 299-304.

40. Palmer, R. H., "200 Years BB (Before Bloom)," Educational Technology, February 1974, pp. 24-25.

41. Pepe, T. J., "Something Special--Enrichment Courses at Very Little Extra Cost," School Management, April 1973, p. 29.

42. Potter, Rachel, "American Education: Some Critical Thoughts," Contemporary Review, Vol. 212, No. 1229 (1968), pp. 300-301.

43. Quignard, J., "Problems of Upper Secondary Education," Comparative Education, Vol. 8, No. 2 (1972), pp. 93-99.

44. Siegel, S., "Nonparametric Statistics for the Behavioral Sciences," McGraw Hill Book Company, Inc., New York, New York, 1956.

45. Simpson, E. J., "The Classification of Educational Objectives, Psychomotor Domain," Illinois Teacher of Home Economics, Winter, 1966-67.

46. Spearman, C., "The Proof and Measurement of Association Between Two Things," American Journal of Psychology, Vol. 15, No. 1 (1904), pp. 72-101.

47. Spearman, C., "Footrule for Measuring Correlation," British Journal of Psychology, Vol. 2, Pt. 1, (1906), pp. 89-108.

48. Stuart, Alan, "The Correlation Between Variate-Values and Ranks in Samples from a Continuous Distribution," The British Journal of Statistical Psychology, Vol. VII, Pt. I, May 1954, pp. 37-43.

49. Tanner, D., "Evaluation and Modification of the Comprehensive Curriculum," High School Journal, Vol. 54, No. 5 (1971), pp. 28-32.

50. Uhl, N. P., "Identifying Institutional Goals: Encouraging Convergence of Opinion Through the Delphi Technique," National Laboratory for Higher Education, Durham, North Carolina, 1971, 86 pages.

51. Unruh, G. G., "Beyond Sputnik," Educational Leadership, Vol. 30 (1973), pp. 587-590.

52. Waldron, J. S., "The Delphi Process: Some Assumptions and Some Realities." Paper presented at the 1971 annual meeting of American Educational Research Association, February 7, 1971, New York, New York.

53. Walen, H. L., "America's High School Dilemma," The National Association of Secondary School Principals Bulletin, Vol. 56, Pt. 2 (1972), pp. 79-85.

54. Wilcoxon, F., "Individual Comparisons by Ranking Methods," Biometrics, Vol. 1, No. 6 (1945), pp. 80-82.

55. Wilhelms, F. T., "Tomorrow's Students," The National Elementary Principal, April 1970, pp. 5-10.

56. Wright, S. A., "Is Separate Education for Vocational Students the Answer?", The Balance Sheet, Vol. 47, May 1966, pp. 398-399.

57. Zar, J. H., "Significance Testing of the Spearman Rank Correlation Coefficient," Journal of the American Statistical Association, Vol. 67, No. 339, Applications Section, pp. 578-580.