THE PREDICTIVE VALIDITY OF THE TEACHER PERCEIVER INTERVIEW ON THE TEACHING PERFORMANCE OF CLASSROOM TEACHERS

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

DANE A. DELLI, B.S., M.ED.

* * * *

The Ohio State University
2000

Dissertation Committee:

Professor I. Phillip Young, Adviser
Professor Philip T.K. Daniel
Professor Robert L. Heneman

Approved by:

[Signature]
Department of Education
ABSTRACT

The employment interview remains a widely used predictor in making employment decisions regarding classroom teachers. This research project investigated the predictive validity of a commercially designed teacher selection instrument formally known as the Teacher Perceiver Interview (T.P.I.), both abbreviated and full versions.

This investigation consisted of two studies. Study I examined 72 teachers hired after being administered the abbreviated version of the T.P.I and Study II examined 124 teachers hired after being administered the full-length version of the T.P.I. In both studies, the predictor variables are the scores attained on the selection instrument and the criterion measures are principal ratings of teacher performance and absenteeism.

Results indicate that, in the two school districts examined, very little relationship existed between scores attained on either version of the Teacher Perceiver Interview and the subsequent job performance measures chosen. Revealed in this investigation is how a lack of internal consistency for this selection instrument brings both reliability and validity issues to bear on both versions of the Teacher Perceiver Interview. Implications of these findings and recommendations for practicing school administrators are discussed.
DEDICATION

This work is dedicated foremost to the people in my life who have helped me become a better person in so many ways. First, to my wife LeeAnn for being both my most avid supporter and my best friend. Thank you for believing in me and for making the personal sacrifices needed to help me reach my goal. You are more responsible for this success than you realize, and I will always be grateful.

I also dedicate this study to a child who has yet to enter our world. May the good fortune that education has given me be but a small fraction of the success you meet throughout your lifetime. Always take time to thank those who have helped you achieve.

To my parents, Sam and Fay Delli, thank you for always encouraging me to follow my dreams and for always allowing me to take chances. With your help and inspiration, I have achieved some very special goals. I will always be grateful for the opportunities you have given me.

I never appropriately thanked a very influential person in my life, Ms. Elizabeth Conser. Your biggest gift was helping me recognize, at a very young age, that all people are different and that I should always find the best in people. This work is dedicated in your loving memory—thanks “Libby.” If Ph.D’s were granted for lives well served, you would have earned yours long ago.
ACKNOWLEDGMENTS

An arduous undertaking such as this is simply not possible without a strong, thoughtful, and insightful mentor. To my Major Professor, I. Phillip Young, I extend my most sincere thanks and appreciation for the countless hours you spent refining my writing skills and building my knowledge base in statistics and research methodology. Not only did you allow me to enter the house of your wisdom, you also lead me to the threshold of my own mind. Your expertise and style of delivery served me exceptionally well.

Although I will be, most likely, unable to pay “back” for what you have given me, I promise to pay “forward” by instilling in my students what you have instilled in me. I am perhaps most grateful for the confidence, encouragement, and wisdom you provided me when I needed it the most. You are an adviser to me in more ways than you know.

To Professor Philip T.K. Daniel, your advocacy and support during my entire doctoral program is clearly a strength that will serve me well as a professor. You have provided me with a foundation in both Education and in Law that strengthens my ability as a researcher focused on helping solve practical problems in schools.

To Professor Robert L. Heneman, your willingness to help me discover the realm of Human Resources Management and your support of my additional research interests
are greatly appreciated. I am very appreciative of the interest you took in me and of the
words of encouragement you often voiced.

Finally, I would like to acknowledge the school districts that agreed to participate
in this study. Administrators in these districts clearly display a strong commitment to
evaluating and improving the human resource function in educational administration.
Specifically, these school districts are addressing both legal and ethical concerns
regarding the testing and employment of classroom teachers – thank you for
participating.
VITA

November 9, 1968.............................. Born – Cleveland, Ohio

1991........................................... B.S. Education
Bowling Green State University
Bowling Green, Ohio

1991-1994................................. English Teacher
Charles F. Brush High School
Lyndhurst, Ohio

1994......................................... M.Ed. Education
John Carroll University
University Heights, Ohio

1994-1995................................. Graduate Research Assistant
The Ohio State University
Columbus, Ohio

1995-1996................................. Principal
Berkshire Middle School
Burton, Ohio

1996-1998................................. Principal
Brooklyn High School
Brooklyn, Ohio

1998-Present.............................. Graduate Research Assistant
The Ohio State University
Columbus, Ohio
PUBLICATIONS


Major Field: Education
# TABLE OF CONTENTS

Abstract ..................................................................................................................................... ii

Dedication .................................................................................................................................. iii

Acknowledgments ........................................................................................................................ iv

Vita ............................................................................................................................................... v

List of Tables ............................................................................................................................... x

List of Figures .............................................................................................................................. xi

Chapters:

1. Introduction .............................................................................................................................. 1

   Problem ................................................................................................................................... 5

   Research Questions .................................................................................................................. 6
   Definition of Terms .................................................................................................................. 7

2. Literature Review .................................................................................................................. 10

   The Organizational Staffing Process ...................................................................................... 13

   An Historical Overview of Selection Research ................................................................... 22
   Macroanalytic Research Perspectives ..................................................................................... 24
   Microanalytic Research Perspectives ..................................................................................... 27
Legal Implications .................................................. 33

Influences of Organizational Selection Research on Teacher Selection Research ......................... 41

Teacher Perceiver Interview:
Historical Background and Advent ........................................ 56

Validation Studies of the Teacher Perceiver Interview ........................................ 62

3. Methodology ............................................................................. 70

Subjects, Study I ........................................................................ 71
Predictor Variables, Study I ................................................. 72
Criterion Variables, Study I ................................................ 73
Procedures, Study I ................................................................. 75
Design and Analysis, Study I .................................................. 76
Subjects, Study II ...................................................................... 77
Predictor Variables, Study II ............................................... 78
Criterion Variables, Study II .................................................. 79
Procedures, Study II ................................................................. 81
Design and Analysis, Study II .................................................. 82

4. Results .................................................................................. 83

Descriptive Data, Study I ......................................................... 84
Descriptive Data, Study II ....................................................... 88
Descriptive Statistics, Study I .................................................. 92
Descriptive Statistics, Study II .................................................. 107

5. Findings, Implications, Limitations ......................................... 127

Findings ................................................................. 130
Implications ................................................................. 141
Limitations ................................................................. 143
Conclusion ................................................................. 144
Appendix A  Glossary of Teacher Perceiver Interview Themes ..................................146

Appendix B  Sample Teacher Perceiver Interview Questions and
            Sample Responses ..................................................................................150

Appendix C  Sample Teacher Rating Scale-Study I .............................................. 154

Appendix D  Sample Teacher Rating Scale-Study II ........................................... 158

List of References ......................................................................................... 163
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive data of school district for study I</td>
<td>85</td>
</tr>
<tr>
<td>2</td>
<td>Demographic characteristics of classroom teachers for study I</td>
<td>87</td>
</tr>
<tr>
<td>3</td>
<td>Descriptive data of school district for study II</td>
<td>89</td>
</tr>
<tr>
<td>4</td>
<td>Demographic characteristics of classroom teachers for study II</td>
<td>91</td>
</tr>
<tr>
<td>5</td>
<td>Descriptive statistics, reliabilities, and correlations for each predicted</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>T.P.I. theme, Study I</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Analysis of items comprising Mission</td>
<td>94</td>
</tr>
<tr>
<td>7</td>
<td>Analysis of items comprising Empathy</td>
<td>95</td>
</tr>
<tr>
<td>8</td>
<td>Analysis of items comprising Rapport Drive</td>
<td>96</td>
</tr>
<tr>
<td>9</td>
<td>Analysis of items comprising Individualized Perception</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>Analysis of items comprising Input Drive</td>
<td>98</td>
</tr>
<tr>
<td>11</td>
<td>Analysis of items comprising Activation</td>
<td>99</td>
</tr>
<tr>
<td>12</td>
<td>Analysis of items comprising Innovation</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>Analysis of items comprising Gestalt</td>
<td>101</td>
</tr>
<tr>
<td>14</td>
<td>Relationship between predictor measures and criterion measures</td>
<td>103</td>
</tr>
<tr>
<td>15</td>
<td>Full regression model (Study I)</td>
<td>106</td>
</tr>
<tr>
<td>16</td>
<td>Theme contribution to regression model (Study I)</td>
<td>106</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Descriptive statistics, reliabilities, and correlations for each</td>
<td>108</td>
<td></td>
</tr>
<tr>
<td>T.P.I. theme, Study II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Mission</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Empathy</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Rapport Drive</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Individualized Perception</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Listening</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Investment</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Input Drive</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Activation</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Objectivity</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Innovation</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Gestalt</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>Analysis of items comprising Focus</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>Relationship between predictor measure and criterion measures</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>Full Regression Model (Study II)</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>Theme contribution to regression model (Study II)</td>
<td>125</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Young’s shared variance model for teacher selection.......................... 45
CHAPTER I

INTRODUCTION

Increased student achievement and the development of standards and benchmarks for learning has been the focus of school improvements efforts since the 1980's (Newman et. al., 1996). Absent from this focus on school improvement has been the impact of the classroom teacher as a significant component to the school improvement process. Webster (1988) maintained, "Much of the output of an educational system seemed logically related to the quality of teachers in that system" (p.245).

A significant amount of research has been conducted supporting the assertion that quality teachers must be recruited, selected, and retained so as to provide quality education to students (Bredeson & Caldwell, 1987; NCTAF, 1997; Trachtenberg, 1973). Despite this evident need to select quality teachers, for the improvement of the educational system, the focus of research efforts and of hiring practices geared toward the selection of quality teachers have changed little over the past several years. Because the selection of quality classroom teachers has been proven essential to the quality of education delivered to students, more emphasis must be placed on the selection of quality teachers.
Every year, school systems experience teacher vacancies due to resignations, retirements, transfers, and new positions (Castetter & Young, 2000). The responsibility for recruiting, selecting, and retaining quality classroom teachers to occupy these teaching vacancies has rested historically with school administrators (Gorton & Schneider, 1991). Merritt (1971) stated,

"The selection of teaching personnel is one of the main functions of educational administrators. The critical nature of this function may be readily seen in the development and implementation of the school educational program. Very often the administrator’s main opportunity to initiate change or strengthen certain functions of the curriculum rests with the decision he [she] makes regarding the selection of teachers with necessary competencies" (p.2).

Selecting classroom teachers, with the necessary competencies to perform the job, is a task so important that practicing administrators must use all available resources to improve the selection system used to hire teachers.

School administrators, seeking to select classroom teachers so as to improve the educational system, must respect also the rights of teacher candidates throughout the employment process. Federal legislation and governmental Acts provide guidelines that afford all individuals an equal opportunity for employment consideration. The challenge facing school administrators is selecting quality classroom teachers while remaining within the acceptable boundaries set forth by law.

The selection of quality classroom teachers begins with the screening decisions of school administrators, based on paper credentials submitted by teacher candidates, during the preinterview stage of the interview process. Teacher candidates, who progress
successfully through the screening stage, are invited then for an opportunity to meet with school administrators, who will make further hiring decisions about teacher applicants during a face to face interview. Although the interview is perhaps the most common employment method used in selecting classroom teachers, research efforts have failed to investigate the impact of selection decisions made during the face to face interview.

Because the employment interview remains such a widely used predictor of future job performance (Dipboye, 1992), more investigation of the employment interview is needed to better understand how improvements can be made within the selection process that would, in turn, increase the reliability and the validity of the interview. Specifically, more attention is needed in evaluating the predictive validity of decisions made as the result of information assessed during an employment interview. Although the validity of the interview has made respectable advances since Wagner's (1949) review of research, more investigation is needed to help raise the psychometric properties of the employment interview.

This study focuses on decisions made by interviewers at the interview stage of the employment process, where administrators interact with job candidates known previously as only "paper people" (Gorman et al., 1978). One type of interview is known as the Teacher Perceiver Interview (T.P.I.). This type of interview, or test according to the Uniform Guidelines (1978), is a commercially designed, structured interview that purports to provide accurate information concerning the future job performance of classroom teachers.
The Teacher Perceiver Interview is composed of sixty open-ended questions designed to allow interviewees an opportunity for self-expression with regard to different job-related issues. This interview is analyzed according to the following twelve themes: (1) Mission, (2) Investment, (3) Focus, (4) Empathy, (5) Rapport Drive, (6) Listening, (7) Objectivity, (8) Individualized Perception, (9) Input Drive, (10) Activation, (11) Gestalt, and (12) Innovation. The Gallup Organization (1997) purports that, “We can better understand and predict job-related behaviors of individuals by using these themes as assumed patterns of positive behavior” (p.1).

Results of the Teacher Perceiver Interview provide information with which administrators make subsequent hiring decisions. The administrative decisions made, as a result of information gathered during an employment interview, can impact, either positively or negatively, the entire educational system. These administrative decisions ultimately impact the intellectual and the developmental growth of all students in the school system.

If, for instance, a highly qualified teacher candidate is extended a job offer then the educational system is afforded an opportunity to improve. Conversely, if an unqualified teacher candidate is extended a job offer, then the improvement of the educational system could be impeded severely. In either case, the educational system is a direct reflection of the teacher quality within the system.

Although a “silver bullet” for always selecting the best and the brightest classroom teachers fails to exist, school administrators must be deliberate in utilizing a selection system that is both reliable and valid. Administrators must remember also that
the selection system employed must follow legally mandated hiring practices. Only by utilizing selection practices that yield valid predictors of job performance can practicing administrators increase their ability to hire high quality classroom teachers.

Problem

Educational administrators are expected to make teacher selection decisions that will affect positively the educational achievement of students and that will help achieve the overall goals of the school system. A teacher selection decision made, during and after the interview stage of the interview process, is perhaps the most significant decision because a decision made to employ a teacher candidate represents generally the end of the interview process. Once a teacher candidate accepts or rejects an employment offer, the educational organization is impacted immediately.

Research supports the need for educators to make wise teacher selection decisions. While this assertion was made early in the development of schools, recent societal events make the assertion for hiring quality teachers even more pronounced. School reform movements and proficiency driven outcomes, coupled with increased legal guidelines, make it incumbent on school administrators to develop a selection system that will ultimately benefit students, while remaining impartial and fair to all teacher applicants.
The problem addressed in this study is the void that exists in the availability of a reliable and valid selection process for identifying outstanding teachers. It is for the reasons stated above that the present study examines the predictive validity of the Teacher Perceiver Interview, both the abridged version and complete version, and on subsequent classroom teaching performance. The following research question are proposed in reference to this study:

Q1: To what extent is there a relationship between a teacher’s Teacher Perceiver Interview Screener (T.P.I.S.) score and the rating of that teacher’s classroom performance by a supervising building-level principal?

Q2: To what extent is there a relationship between a teacher’s Teacher Perceiver Interview Screener (T.P.I.S) score and the number of days the teacher missed work?

Q3: To what extent is there a relationship between a teacher’s Teacher Perceiver Interview (T.P.I.) score and the rating of a teacher’s classroom performance by a supervising building-level principal?

Q4: To what extent is there a relationship between a teacher’s Teacher Perceiver Interview (T.P.I) score and the number of days the teacher missed work?
Definition of Terms

1. **Bona fide occupational qualification** – a statutory provision that permits discriminatory practices in employment where the discrimination against the protected class is reasonably necessary to the normal operation of the particular organization.

2. **Concentration statistics** – an analysis used to illustrate that significant differences exist between and among protected class groups, as distributed throughout levels of an organization or as compared with other organizations.

3. **Disparate impact** – a term describing a type of discrimination in which the actions of the employer, consciously or unconsciously, lead to discrimination against a member of members of a protected group. Not only the act of discrimination, but also the outcome of the discrimination is considered.

4. **Disparate treatment** – a type of discrimination in which the intentions of the employer lead to discrimination, such as discriminatory words, phrases, or signs.
5. **Flow statistics** – an analysis used for comparing employee selection rates for persons of protected class status against the selection rates for persons not classified as members of a protected class.

6. **Macroanalysis** – the study of the selection process focusing on the relationship between an independent variable(s) (predictor) and a dependent variable (criterion). Researchers used macroanalysis to examine the employment interview as a single step process.

7. **Microanalysis** – the study of the selection process focusing on identifying factors that do not directly affect dependent variables under examination, but still affect the independent variables used in decision making. This research perspective examines the employment interview in various stages, rather than as a single event.

8. **Prima facia evidence** – evidence of discrimination that does not require further support to establish a valid case.

9. **Stock statistics** – an analysis used for illustrating the percentage of protected class employees who are available in either the general population or general labor market in a geographic region.
10. **Teacher Perceiver Interview (T.P.I.)** – a commercially designed, structured interviewing system purporting to gather relevant information about a job candidate and to identify a person’s potential for outstanding job performance.

11. **Teacher Perceiver Interview Screener (T.P.I.S.)**– an abbreviated version of the original, full-length, Teacher Perceiver Interview designed as a “time saver.” This version of the T.P.I. measures ten of the twelve T.P.I. themes. Themes “Listening” and “Focus” are absent from the T.P.I.S.

12. **Teacher Themes** – twelve reoccurring patterns of thought, feeling, and behavior purporting to represent predictable aspects of an individual’s personality and an individual’s behavior.
CHAPTER 2

LITERATURE REVIEW

The education of children is the central purpose of any school, and the single most important resource in producing a quality education is the classroom teacher. Stakeholder groups such as administrators, teachers, parents, and board members have a vested interest in making wise teacher selection decisions. It is possible also that each stakeholder group maintains fundamentally different philosophical beliefs as to what constitutes a good classroom teacher and what constitutes good classroom instruction.

One stakeholder group may, for instance, have firm beliefs that the primary responsibility of classroom teachers should be to help students attain predetermined achievement standards. Conversely, another stakeholder group may have firm beliefs that the primary responsibility of classroom teachers should be to enhance individual student learning, independent of predetermined achievement standards. Regardless of any differing philosophical beliefs about the primary responsibility of the classroom teacher, it is the selection of qualified teachers in a school district, not procedures, or handbooks, or regulations, or curriculum guides, that is the most important tool a district has to work with when seeking to improve (Gorton & Schneider, 1991).
The issue of school improvement began to receive frequent attention during the 1980’s when several reports were commissioned to investigate the status of American education and to make recommendations for educational improvement (see Boyer, 1983; Goodlad, 1984; National Commission on Excellence in Education, 1983). Among several recommendations for improvement was increased effectiveness of teacher instruction and increased student achievement. One of the most fundamental ways educational organizations can ensure that students are achieving at acceptable levels is to employ competent, motivated, and dedicated classroom teachers.

Unfortunately, the increased employment of qualified classroom teachers has improved little since the instructional improvement mandates of the 1980’s. According to the National Commission on Teaching and America’s Future (NCTAF), which published a report entitled, “What Matters Most: Teaching for America’s Future” (1996), there exists a very disconnected approach as to how teachers are recruited, selected, inducted, and rewarded. The NCTAF report indicated that, “In too many places, recruitment and hiring practices are out of sync with new student standards” (p.3).

Cited in the NCTAF report were studies revealing that differences in teacher qualification and experience “can account for most of the difference in student achievement – even when factors such as home background are taken into consideration” (p.4). Compelling testimony, for the lack of improved student achievement, is perhaps illustrated best by recognizing that in 1991, “over 12% of new hires entered the classroom without any formal training at all” (p.3). Furthermore, in some schools,
“students have less than a 50% chance of getting a science or mathematics teacher who
holds a license and a degree in the field they teach” (p.4).

Indeed, the selection of unqualified teachers can thwart the goal attainment of a
school district and can stunt the educational growth of students in a district. Graces
(1932) made a profound statement that remains very true today: “Wise teacher selection
is the best means of improving the system, and the greatest lack of economy exists
whenever teachers have been poorly chosen, or are inadequately adapted to their
profession” (p.191). Bolton (1969) supported further the importance of hiring quality
teachers by stating,

“So crucial is the selection of a teacher to the quality of the educational program,
that it seems obvious that this decision should be made with the utmost certainty
regarding it’s utility. Yet such decisions are frequently intuitive and arbitrary, and
despite a certain amount of theory development, a lack of empirical data has left
the teacher selection process a highly subjective one” (p.329).

Because educational administrators, who are charged with hiring all school
employees, have experienced difficulty attracting and employing well qualified
classroom teachers, a fundamental understanding of the staffing process is essential.
Only by understanding ways to improve the staffing process can educational
administrators hope to make wiser teacher selections that, in turn, improve classroom
instruction and improve student achievement. A conceptualization of how educational
administrators should view the staffing process is best illustrated by using an
organizational staffing process as a model.
The Organizational Staffing Process

Staffing, as defined by Heneman, et al. (1997), is “the mutual process by which the individual and the organization become matched to form the employment relationship” (p.4). This process is referred to as “mutual” because individuals seek positions within organizations just as employers seek individuals to occupy positions. It is through this mutual decision making process that a person and a job fit is obtained.

From the viewpoint of the employer, a person and a job fit is central to the staffing process. In assessing a person and a job fit, employers select specific job predictors, using a job analysis, that are assumed to have a correlation with future job performance. Heneman, et al. (1997), referred to job analysis as the “process of studying jobs in order to gather, analyze, synthesize, and report information about job content” (p.160). A job analysis, which represents the foundation of the staffing process, consists of two fundamental components: (1) job requirements and (2) job rewards.

One of the fundamental components to job analysis is job requirements. Job requirements are comprised of both job tasks and knowledge, skill, ability, and other characteristics (KSAO’s) needed to perform designated tasks. Job tasks are used to constitute job descriptions, while KSAO’s are used to constitute job specifications.

Once job requirements have been established, job requirements are linked with job rewards. Job rewards, which are associated with every job, are categorized generally as extrinsic rewards and intrinsic rewards (Heneman, et al., 1997). Extrinsic rewards
include factors such as salary and fringe benefits, while intrinsic rewards include factors such as job satisfaction and intellectual stimulation.

After a job analysis, consisting of job requirements and job rewards, has been completed, the staffing process begins. The staffing process is comprised of three core components: (1) recruitment, (2) selection, and (3) employment. Each of these component parts represents a distinct phase within the staffing process.

The initial phase of the staffing process is recruitment. Recruitment practices are designed to assist organizational representatives in identifying and attracting individuals as prospective applicants. Organizational representatives, who recruit applicants, must, for example, be cognizant of where applicants will seek employment opportunities.

One rather consistent research finding indicates that individuals seeking employment opportunities tend to rely on non-traditional recruitment sources. A non-traditional recruitment source, which has become a popular vehicle to recruit future employees, is school districts (Bainbridge & Sundre, 1991). Traditional recruitment sources, by which organizations seek employees, include formal advertisements, employee referrals, college placement offices, and employment agencies.

Just as employers identify where prospective applicants will seek employment opportunities, employers must understand also how the recruitment phase includes mutual decision making between the applicant and the employer. Recruiters form impressions, for example, about applicants interested in gaining employment. The recruiter is confronted also with making decisions as to the possible “fit” of the candidate within the organization.
Applicants are making decisions also about whether an organization or a particular job is appealing. Factors such as job information and behaviors of the recruiter influence the impressions of applicants. Research has shown that both well articulated job information and observable behaviors of the recruiter have a significant impact on applicant decision making (Dipboye, 1992).

Because the behaviors of the recruiter impact the decisions of the applicant, deliberate selection and training of recruiters is important in the planning and implementation of recruitment activities. One study revealed that an ideal recruiter would possess the following characteristics: strong interpersonal skills, knowledge of the organization, a focus on career-related issues, and enthusiasm about the organization and job candidates (Rynes & Boudreau, 1986). After careful consideration has been given to the planning and the implementation of recruitment activities, organizational representatives can begin the second phase of the staffing process: the selection phase.

The selection phase of the staffing process begins with assessment of and with evaluation of the applicant pool created during the recruitment phase. During the selection phase, an interview is employed to help determine the person and the job fit through assessment and evaluation. Two researchers maintained that, just as staffing is a process, the selection interview is a three stage process as well (Dipboye, 1992; Macan & Dipboye, 1990).

First, paper credentials of applicants are screened during the preinterview stage of the selection interview, consisting of three stages, is a process. Second, both the applicant and the employer interact face to face during the actual interview stage of the
selection interview process. Third, employers make final hiring decisions during the employment stage of the selection interview process.

During the preinterview stage, employers first screen paper credentials provided by applicants. It is during this stage of the interview process when employers are first “introduced” to applicants. Gorman, et al., (1978) referred to applicants during the preinterview stage as “paper people.” The paper credentials submitted by applicants for review by the employer include relevant documents such as application forms, personal resumes, reference letters, and academic transcripts.

The most commonly reviewed paper credential is the application form. An application form is an advantageous means of collecting data about applicants because the employer can determine what information the applicant will provide and because the employer can use the document to verify information presented on a resume. By requiring predictor specific information from applicants, employers can judge better future employee performance on the basis of resume information.

Another paper credential reviewed frequently by employers is the applicant resume. Resumes reference typically work experiences, educational attainments, personal qualifications, job skills, and related accomplishments. Young and Ryerson (1986) warned, however, that resumes should be used with caution because some resumes received by employers “may have been prepared by professional firms while other resumes may have been developed by individual job candidates”(p.8). Although little research has been conducted on the resume format, Helwig (1985) indicated that a
traditional, conservative resume format yields more favorable responses from
interviewers than non-traditional liberal resumes.

Still another type of information provided by applicants is reference letters. Although reference letters are used frequently, these types of letters fail generally to yield much useful information on applicants. Muchinsky (1979) and Young (1986) established that reference letters contain often very general information that is overstated and very positive.

Yet another paper credential reviewed by interviewers is an applicant transcript. Transcripts contain predictive data such as academic credentials and test scores that interviewers can utilize as they seek to establish a fit between the person and the job. Also, transcripts represent a means by which employers can secure official documentation as to degree status and to program of studies for an applicant.

During the preinterview stage, employers attempt to reduce further the applicant pool. By utilizing substantive assessment measures as predictors of future job performance, employers can create a group of finalists for employment consideration. In addition to the previously mentioned paper credential assessment measures, employers utilize tests during the preinterview stage.

One type of preinterview test is a personality inventory that is used to assess the general disposition of a job candidate. Another type of test, such as an ability inventory, includes cognitive, psychomotor, and sensory perception measures. Still another type of test is a physical dexterity test used to help determine the physical fitness and physical ability of a job candidate (Dipboye, 1992).
Based on the information gathered and screened during the preinterview stage, interviewers make decisions about which applicants will be invited to the next stage of the interview process and which applicants will be excluded from further employment consideration. Those applicants who are invited for further consideration make decisions also about whether to remain a potential employee or to self-select out of the interview phase. The decision of an applicant to self-select out of the selection process impacts the organization by leaving the employer with a smaller applicant pool from which to make selection decisions.

Applicants, who decide to continue in the selection process, will be affected also by the impressions formed by the interviewers during the preinterview stage. Research has shown that interviewers are influenced more by negative information than by positive information and that interviewers fail to place enough weight on positive information. As a result, the information provided by the applicant during the preinterview stage is significant (Bolster & Springbett, 1961; Hollman, 1972; London & Hakel, 1974).

Candidates, who have made successful progress through all of the above-mentioned preinterview activities, proceed to the second stage of the interview phase. The second stage of the interview phase is the face to face interview, between the employer and the applicant; which, has maintained an important and frequently utilized role as a predictor in the selection process. In fact, one researcher predicted that the interview will continue to be the most commonly utilized predictor of job performance (Dipboye, 1992).
Dipboye (1992) defined a selection interview as "a dialogue initiated by one or more persons to gather information and evaluate the qualifications of an applicant for employment" (p.3). During the selection interview, applicants known previously only as "paper people" are given now an opportunity to interact with the employer. It is at the interview stage where the initial impressions of the applicant, formed during the preinterview stage, have been shown to influence employer decision making and to influence the outcome of the interview (Macan & Dipboye, 1990).

Because information processing and decision making are key components to the employment interview, a great deal of research has been focused on the cognitive processes of the interviewer. Dipboye (1992) pointed out research that showed "people hold firm views of the personality traits possessed by persons in occupations largely based upon stereotypes" (p.18). It is often through these stereotypes that the judgment of the interviewer is influenced most.

One inherent problem with judgment stereotypes is that employment decisions are sometimes made based on irrelevant applicant information. Furthermore, judgment stereotypes based on irrelevant data impacts the decision of the interviewer very early in the interview. Some categories of irrelevant applicant information include sex, race, and socioeconomic level (Dipboye, 1992).

Applicants are often categorized on particular traits or dispositions within cognitive categories. For example, a five-factor personality scale, developed by Hakel (1974), outlines the following five personality categories: (1) extroversion, (2) agreeableness, (3) conscientiousness, (4) emotional stability, and (5) culture. These
cognitive categories appear to be hierarchical also in that people begin by holding general beliefs about applicants and, as more information is presented, the more defined the cognitive category becomes and the more pronounced the interviewer judgment.

Judgments made by interviewers may sometimes be influenced by the level of training and of experience possessed by the interviewer. Fiske and Kinder (1981) maintained that interviewers with more training and more experience, hold more complex schemas than interviewers with less training or less experience. Another study examined the effects of using college students as interviewers rather than actual personnel managers as interviewers and found that evaluations by student interviewers lack generalizability to the job setting (Singer & Sewell, 1989).

The final stage of the selection interview is the employment stage. During the employment stage of the interview, a decision is made to offer a position to one applicant and to deny an offer to another applicant. Some of the methods used to help employers decide which finalists will be offered jobs, and which finalists will be denied jobs, include random selection, ranking, and grouping.

One method employed in deciding which finalist will be offered a position within the organization is the random selection method. Random selection refers simply to one finalist being chosen indiscriminately from a group of finalists. This random selection method can be used fairly when there is more than one applicant who is worthy of being offered a position and when each candidate has an equal chance of being offered a position.
Still another finalist selection method, that remains popular among employers, is the ranking method. When using the ranking method, employers rank order desirable finalists from most desirable to least desirable. Finalists are offered positions then strictly in the order in which finalists are ranked.

Yet another finalist selection method used by employers in making final selection decisions is grouping. The grouping method is a vehicle employers use to band finalists together into rank-ordered categories. Finalists are grouped simply according to whether these prospective employees are “top choices,” “acceptable choices,” or “last resorts,” and finalists are then ranked within the respective groupings according to the desirability of the employer.

Irrespective of the final selection method used by employers, ensuring the best possible person and the best possible job fit is the cornerstone of a good selection decision. The selection of individuals, who lack the necessary knowledge, skills, abilities, and other characteristics needed to perform a particular job, will impact adversely the goals of the organization. Heneman and Heneman (1994) acknowledged the importance of attracting, hiring, and retaining the most qualified individuals because it is the human resources of any organization that play a fundamental role in determining the success or failure of the organization.

The organizational staffing process, consisting of a recruitment stage, a selection stage, and an employment stage, is the product of research efforts spanning many years and involving many different researchers. Research addressing the selection stage of the staffing process is specifically the focus of the present study. An historical overview of
selection research, from both a macroanalytic and a microanalytic perspective, will establish a context for analyzing further selection research.

An Historical Overview of Selection Research

Two Research Perspectives: Macroanalysis and Microanalysis

The selection interview has been the topic of research efforts for over 80 years. As a result, many significant contributions have been made to the analysis and to the development of the selection interview (Arvey & Campion, 1982; Campion et. al., 1997; Harris, 1989; Mayfield, 1964; McDaniel et al, 1994; Schmitt, 1976; Ulrich & Trumbo, 1965; Wagner, 1949; Wright, 1969). As research findings continue to build on this current body of knowledge, it becomes necessary for both practitioners and theorists to recognize and to understand two fundamental research approaches to the selection interview: (1) macroanalysis and (2) microanalysis.

Most early research on the selection interview originated from a macroanalytic perspective. Young and Ryerson (1986) maintained that the macroanalytic approach to the selection interview has two major goals. One of the goals of macroanalytic research is to assess the relationship between an actual criterion measure representing job performance and a specific predictor variable purporting to measure job performance. The second goal of macroanalytic research is to maximize the amount of variance in job performance that can be accounted for by a predictor variable.
Research proponents of the macroanalytic approach sought to assess the reliability and validity of the employment interview and to improve the interview as a predictive selection instrument. Young (1985) maintained that macroanalytic studies included typically a composite interview score related to employability indices with a measure of job performance and were based on field data. Because macroanalytic research failed to produce optimistic results, Mayfield and Carlson (1966) conducted research that triggered a philosophical and a methodological shift from a macroanalytic perspective to a microanalytic perspective.

Microanalytic research efforts focused on the decision making process and on factors that influence the decision making process, rather than the direct relationship between a predictor variable and criterion obtained from evaluating job performance. Young and Ryerson (1986) stated, “Microanalytic research has sought to identify those variables that are unrelated to performance and yet influence the selection decisions made by school administrators” (p.26). In determining those variables unrelated to performance, microanalytic investigators employed experimental research such as interview simulations and resume simulations of “paper people” as hypothetical candidates (Young, 1985).

Now that a brief overview of both the macroanalytic perspective and the microanalytic perspective is provided, a review of selection interview literature, using these research perspectives will be presented. The application of the macroanalytic and of the microanalytic perspectives to the employment interview provides the foundation from which all selection interview research has evolved. Only though understanding
where selection research originated, can an understanding be obtained of where selection research should be now directed.

**Macroanalytic perspectives**

Early research in personnel selection focused on the selection interview as a predictor of job performance. This early research was considered macroanalytic (Webster, 1964), in that, investigators concentrated on the relationship between a specific predictor of job performance and an actual criterion measure of job performance. Researchers, who utilized the macroanalytic approach, focused on reliability and validity as a vehicle for improving the interview as a selection devise (Young, 1985).

Wagner (1949) was the first researcher to publish a comprehensive review of the literature on the employment interview. His review was representative of an organizational perspective, and the review focused on the selection interview as it related to the overall ability of the applicant as a predictor of job performance. This investigator was interested most in the potential of the interview as a valid means of predicting success on job performance.

The review conducted by Wagner revealed clearly a lack of reliability and a lack of validity for the selection interview as a predictor of future job performance. Wagner began his review by citing one of the earliest investigations of the interview published by Binet (1911) who reported virtually zero inter-rater reliability among interview assessed data collected from three teachers who were interviewing the same five children.
In producing his review, Wagner was able to locate 106 articles investigating some facet of the employment interview. Only 25 articles, however, yielded any empirical data. Furthermore, of the 22 validity coefficients summarized by Wagner, there was a range from .09 to .94 with a median coefficient of .27.

Wagner cited a study by Wonderloric (1942) suggesting that the interview is a poor choice for measuring mental ability or job skill and that these types of measurement are suited better for scientific other assessment methods, such as testing. Also cited by Wagner was Kelley (1929) who indicated that certain kinds of appraisal could be made in no other manner except in a situation involving continuous social interaction. Based on his findings, Wagner recommended that the interview may prove useful in three situations: first, when a rough screening procedure is needed; second, where the number of applicants is too small to warrant other, more involved procedures, and third; where certain traits may be evaluated more accurately by an interview than by any other means (Wagner, 1949).

Also advocated by Wagner was the standardization of the employment interview. Some of the advantages of standardization included that the interviewer works from a job description, knows what questions to ask, is trained in interviewing techniques, and has a “yardstick” by which to interpret and to evaluate information gleaned from the applicant. It is not surprising then, that, one of Wagner’s concluding remarks was “… the whole matter of the interview as a method of selection and placement remains in a state of confusion” (p.33).
Wright (1969), who was a proponent of the macroanalytic approach to selection research, conducted a survey of selection studies since Wagner's (1949) review. In building a case for macroanalytic approaches to research, Wright used microanalytic studies. Wright, for example, cited Webster (1961) who conducted a study on the decision making process during the employment interview. Basic findings included that interviewers develop stereotypes and match interviewees with those stereotypes; negative information is more influential than positive information; and, decisions made by the interviewer vary often when information is provided “piecemeal” rather than simultaneously.

In proposing general conclusions, Wright (1969) stated,

“Much microanalytic research is so contrived as to approach a level of patent sterility that precludes replication or extension of the line of investigation. While one can understand the call for more microanalytic research, it is possible to deplore the preciousness of this research on the selection interview” (p.409).

Wright claimed, in his analysis, that the structured interview demonstrated rather consistent reliability, but that research efforts should continue to investigate the validity of the structured interview.

Another advocate of macroanalytic research was Schmitt (1976). Based on his findings, he argued that microanalytic research was not sufficient due largely to a lack of integration. Schmitt stated, “The generalizability of these artificial and microanalytic studies to real employment situations remains unestablished” (Schmitt, 1976, p.79).

In summarizing his more significant findings, Schmitt maintained that (1) use of a structured interview guide increased inter-rater reliability, (2) interviewers having more job information about a particular position increased inter-rater reliability, and (3) non-
verbal cues account for some of the differences between interviewer ratings (Hakel, 1974). Another finding was that interviewers weigh negative information more heavily than positive information and that interviewers make final decisions quite early in the interview (Springbett, 1958). Schmitt (1976) cited also Farr (1973) who found that the order in which positive and negative information is presented has an impact on the decision making of the interviewer.

Microanalytic Perspectives

Despite the frequent use of the employment interview, macroanalytic studies of the interview yielded disappointingly low reliabilities and low validities (Wagner, 1949). Only after research conducted by Mayfield and Carlson (1966), was there a movement from a macroanalytic perspective to a microanalytic perspective. Researchers use the microanalytic perspective to examine how the decision making process affects selection decisions.

After reviewing over 300 articles involving the selection interview, Mayfield (1964) concluded that very little information had been gleaned regarding the effectiveness of the employment interview in predicting job performance. Mayfield stated, “Our knowledge of the selection interview is only a little more advanced than it was when Wagner reviewed selection literature in 1949” (Mayfield, 1964, p.248). His criticisms of the macroanalytic perspective centered on what he believed was a lack of substantive support for the reliability and for the validity of a selection interview.
Mayfield's review, therefore, resulted in several recommendations for improving the reliability and the validity of a selection interview. In supporting the use of the microanalytic approach, Mayfield cited Webster (1959) who investigated how interviewers reach final selection decisions. As a result of Webster's findings, Mayfield asserted that research efforts should be directed toward examining how selection decisions are reached, within the selection context, rather than on the specific outcomes of the interview.

Mayfield's review and analysis of published research on the employment interview, as a result, produced 15 fundamental conclusions (see pp.248-254). Perhaps the most notable conclusion Mayfield drew was that the interview could be best proven reliable by dividing the interview into units. Among some of his other conclusions were that differences existed between interviewer perception and weighting of positive and negative information, and that decisions appeared to be made early in the interview.

Another significant recommendation made by Mayfield, which supported Wagner's earlier finding, was the continued use of structured interviews. While Mayfield asserted that a structured interview increased reliability, he raised questions as to the real effects of structuring the interview. He claimed that it was difficult to ascertain if a structured interview form, which is imposed on the rater, was responsible for the higher reliability, as reported by Mass (1963).

Mayfield began to forge also a new direction for selection research by suggesting the use of multiple interviewers during the employment interview. He asserted that "where tests have a reasonably high relationship with success, the use of the typical
individual interview is of little, if any, value” (p.244). This investigator suggested further, though, that this assertion may not necessarily be true when more than one interviewer is utilized and that the use of a “board” interview may be a promising interview method.

Soon after Mayfield’s review, Ulrich and Trumbo (1965) validated some of Mayfield’s conclusions concerning microanalytic approaches to investigating the employment interview. These researchers found little evidence regarding the validity of the interview and concluded that the interview was best suited to assess the personal attributes of a job candidate. Ulrich and Trumbo did, however, support Wagner by indicating at least one way to increase the validity of the employment interview.

The most significant way these researchers found to increase the predictive validity of the employment interview was to use test data and other ancillary sources such as personal resumes, job applications, and employment tests. Ulrich and Trumbo indicated also that validity increases incrementally as ancillary data is added to the interview. These findings present compelling testimony that the selection interview is indeed, part of a process.

As microanalytic research progressed, Arvey and Campion (1982) reaffirmed that positive reliability and positive validity results could be obtained by using ancillary data and that “Test usage may become viable as an alternative or supplement to interviews for selection purposes” (p.315). While affirming the use of ancillary data, the investigators discovered also the presence of interviewer bias within the selection process. As a result,
the researchers suggested future investigations with respect to the impact of variables such as race, gender, and age.

Although Arvey and Campion subscribed to the belief that microanalysis was needed in evaluating race, gender, and age effects, these researchers advised also using other methods of analysis as well. Arvey and Campion, for example, encouraged more research using the face to face interview. These researchers suggested using person-perception theories to understand better the employment interview by stating “the evaluation judgments made by interviewers are surely a function of attributions” (Arvey & Campion, 1982, p.313).

Subsequent selection research conducted by Harris (1989) revealed a positive trend regarding both the reliability and the validity of the employment interview. He noted specifically findings that indicated the presence of modest validity for the interview as a selection device. Many of Harris’ findings supported earlier studies, while some of his findings contradicted earlier studies.

In supporting earlier research findings, Harris (1989) found that structured interviews appeared to yield stronger correlation coefficients, with performance criterion, than unstructured interviews. He found also that interviews based on a prepared job analysis appeared more valid than interviews lacking a prepared a job analysis. Harris cited also studies that indicated the interview appeared to be predictive of job performance, regardless of the type of interview employed.

Contrary to earlier research findings of Arvey and Campion (1982), regarding gender bias, Harris (1989) maintained that gender had little or no effect on interviewer
ratings. He noted further, however, that his findings regarding race bias on interviewer ratings were inconsistent. This investigator recommended further research on potential racial bias within the selection process.

Harris purported that, because the interview is at least a moderately valid predictor of job performance, future research efforts should focus on why the interview works as a predictor. He recommended, for example, researching why structured interviews are a more valid predictor of job performance than unstructured interviews. Although predictive validity has been shown to exist, Harris maintained that “There appears much that can still be learned from a microanalytic approach” (Harris, 1989, p.720).

More recently, McDaniel et al., (1994) concluded that the validity of the employment interview is higher for structured interviews than for unstructured interviews. Furthermore, these researchers asserted that validity was similar for job performance criteria and for training performance criteria. Even with varying degrees of validity, depending on structure, the employment interview remains the most widely used selection process for staffing organizations.

While structured interviewing has demonstrated increased validity, utilizing multiple interviewers has shown to increase reliability. The team approach to interviewing has been suggested by many researchers as a promising selection method (Arvey & Campion, 1982; Mayfield, 1964; Pursell, Campion & Gaylord, 1980; Roth & Campion, 1992; Ulrich & Trumbo, 1965; Weisner & Cronshaw, 1988). Unfortunately,
since this recommendation was made, little empirical research has been published as to what methodological effects are present when more than one interviewer is utilized.

Weston and Warmke (1988) reported that “at least 20 percent of job seekers will face a group interview somewhere in their job search” (p. 109), and the percentage could be higher for those in public sector positions or academia. A “group” interview, which is synonymous also with “panel” interview and “board” interview, is understood to operate when more than one interviewer is being utilized. Panel interviews are utilized often for civil service selection procedures because the panel “reduces the impact of idiosyncratic biases that single interviewers might introduce” (Campion et al., 1988). Furthermore, Weisner & Cronshaw (1988) asserted that panel interviews should be more reliable than individual interviews because multiple independent ratings are collapsed into a single composite rating.

Another advantage to panel interviews is that the structure of the panel is flexible in that the size of the panel, the composition of the team members, the length of time, as well as with whom the final selection decision will lie, are all variables that can be controlled. Weston and Warmke (1988) maintained that panel interviews are likely also to produce better decisions because the group consensus process mitigates often the biases that interviewers might otherwise display during an individual interview. One clear advantage of the panel interview is, therefore, that an inaccurate perception of one panel interviewer can be corrected by another panel interviewer, who can provide additional insight or additional explanation for an interviewee response.
Regardless of whether an individual interview or a group interview is utilized, fairness in recruiting, selecting, and employing individuals is critically important. Establishing and maintaining a fair selection process is mandated through new and revised federal and state legislation and governmental acts that prohibit discriminatory employment practices and affords all individuals an equal opportunity for employment. Before considering methodologies employed in making selection decisions, an understanding of legal mandates that drive the selection process is essential.

**Legal Implications**

Employment laws and regulations play a major role in the selection process, and these laws and regulations are designed, in part, to reduce or to limit the potential for discrimination of an organizational representative in the employment process. Only if an employer can demonstrate the necessity of a bona fide occupational qualification, can an employer impose job qualifications that would otherwise be considered illegal. There are essentially four major legislative acts, created to ensure fairness, which have a profound impact on the selection process. These major legislative acts are the following: (1) Title VII of the Civil Rights Act of 1964, (2) Age Discrimination in Employment Act of 1967, (3) Section 504 of the Rehabilitation Act of 1973, and (4) Americans with Disabilities Act (ADA) of 1990.

One very significant governmental mandate is Title VII of the Civil Rights Act of 1964, which makes it unlawful specifically for an employer to discriminate against any individual on the basis of race, color, religion, sex, or national origin, as it relates to any
of their employment rights. This Act prohibits further classifying or segregating employees or applicants because of race, color, sex, or national origin. The 1972 amendment broadened the original industrial scope of this law to include governmental institutions and educational institutions.

Another significant employment law is the Age Discrimination in Employment Act (ADEA) of 1967. This Act prohibits specifically discrimination against any person 40 years of age and older with respect to terms and conditions of employment or compensation. ADEA prohibits further classifying or segregating employees or applicants on the basis of age.

Still another legislative employment act is the Rehabilitation Act of 1973 (Section 504), which prohibits discrimination against persons who possess, have possessed, or are believed to possess "a physical or mental impairment that substantially limits one or more major life activities" (Bindboye, 1992, p.188). Although this act fails to require the hiring of unqualified handicapped workers, the Rehabilitation Act does require plans for handicapped workers to include steps to accommodate workers with disabilities so they may perform jobs they might otherwise be unable to perform (Arvey & Faley, 1992). Section 504 of the act, however, applies to only to those individuals who are "otherwise qualified." A qualified individual with a disability is "an individual who, with or without, reasonable accommodation, can perform the essential functions of the employment position that such individual hold or desires" (Arvey & Faley, 1992, p. 64).

The protection afforded to job applicants with disabilities, in the Rehabilitation Act of 1973, was expanded further in 1990 with the Americans with Disabilities Act
(ADA). A significant component to ADA is Section 102, which contains the following provision:

"No covered entity shall discriminate against a qualified individual with a disability because of the disability of such individual in regard to job application procedures, the hiring, advancement, or discharge of employees, employee compensation, job training, and other terms, conditions, and privileges of employment."

This provision is significant because it bans discrimination explicitly in any facet of the employment context except when the employer deems certain functions essential to a particular job. An employer can substantiate what functions of a particular job are essential by preparing a written description of a job before advertising or before interviewing candidates for the job.

The ADA prohibits further pre-employment inquiries regarding specific disabilities, but the act does permit the employer to ask whether a candidate can perform any or all of the job functions established previously as "essential." There exists, however, some debate as to the full implications of this act for interviewing and for other selection procedures (Dipboye, 1992). Employers are encouraged, therefore, to identify and to explain carefully those functions essential for a particular position.

In identifying any job function essential to a particular position, employers must adhere to the above mentioned equal opportunity legislation. Employers must avoid making employment decisions based on protected class characteristics unless those characteristics represent clearly a bona fide occupational qualification (BFOQ). A BFOQ is defined as an employment qualification "reasonably necessary" for the normal operation of the organization (Heneman et al., 1997). In some instances, although quite
rare, certain occupations may necessitate an employee to be either a particular sex or a particular age, but there are no occupations that necessitate an employee to be a particular race or color (Breaugh, 1992).

Careful attention must be given to all federal guidelines throughout the entire selection process because “the interview process is particularly vulnerable to subjective biases, prejudices, and stereotypes on the part of interviewers” (Arvey & Faley, 1992, p.55). One particular protected class applicant characteristic shown to influence selection decisions of employers is applicant age (Young & Schmidt, 1988). Employers, who allow protected class characteristics to influence selection decisions, will be accused likely of unlawful discrimination.

For a protected class applicant to assert a discrimination claim, the applicant must demonstrate merely *prima facia* evidence that unfair discrimination has occurred. According to Dipboye (1992), *prima facia* means that a case for discrimination is established until the party, accused of discriminating rebuts the claim. Two doctrines exist for illustrating *prima facia* evidence: (1) disparate treatment and (2) disparate impact (Young & Ryerson, 1986).

Disparate treatment is the first doctrine, which refers to intentional employer discrimination based on the deliberate exclusion of protected class applicants from employment consideration. The doctrine of disparate impact is the second doctrine, which refers to the effect that current employment practices have yielded currently or have yielded over time. Once a protected class person establishes that an employer has
made employment decisions based unfairly on protected class status, the burden of proof to refute the claim shifts to the employer (Castetter & Young, 2000).

The basic assertion of intentional discrimination, on behalf of an employer, had been difficult for plaintiffs to establish until *McDonnel Douglas Corp. v. Green* (1973). In this case, The Supreme Court created a four-step process that plaintiffs could follow in establishing a prima facia case for disparate treatment. These steps provided an avenue to demonstrate that plaintiffs:

(i) belong to a protected classification; (ii) they applied and were qualified for a job for which the employer was seeking applicants; (iii) despite their qualifications, they were rejected, and (iv) after their rejection, the position remained open and the employer continued to seek applicants from persons of complainants’ qualifications” (*McDonnel Douglas Corp. v. Green*, 1973 as cited in Arvey & Faley, 1992 p.79).

While disparate treatment involves the intentions of the employer, disparate impact involves the actions of the employer. Plaintiffs who assert a claim of disparate impact must present prima facia evidence that the selection process affects adversely members of a protected class. Prima facia evidence used to illustrate disparate impact includes three types of statistical analysis: (1) flow statistics, (2) stock statistics, and (3) concentration statistics (Castetter & Young, 2000).

Flow statistics are used when comparing selection rates for persons of protected class status against the selection rates for persons failing to be classified as having protected class status. In *Albemarle Paper Co. v. Moody* (1975), the plaintiffs established successfully a prima facia case of disparate impact by using flow statistics to illustrate “significant disparity” between the percentage of black applicants, as opposed to the
percentage of white applicants, who passed a paper and pencil employment test.

“Significant disparity” was clarified later when the Uniform Guidelines (1978) outlined that “A selection rate for any racial, ethnic, or sex subgroup which is less than four-fifths (4/5) (or 80 percent) of the rate for the group with the highest rate will generally be regarded . . . as evidence of adverse impact” (p.38297).

As discussed by Daniel and Timken (1999), the applicability of demonstrating the disparate impact doctrine, within the educational arena, continues to be debated among legal scholars. Specifically, in institutions of higher education, strong opinion exists as to the appropriateness of affirmative action practices in college admissions. Court rulings such as Hopwood v. University of Texas (1996) and Bakke v. University of California, Davis, illustrate the distinct conflicting opinion on using race in making college admission determinations. These differing opinions are testimony to the apparent difficulties in demonstrating disparate impact within the educational community of the public sector.

From an organizational perspective, the use of flow statistics is sometimes problematic in that few protected class applicants may have participated in selection activities due to their under representation in the surrounding geographic area. As a result, organizations will utilize stock statistics to determine the presence of or absence of discriminatory employment practices. Stock statistics represent the percentage of protected class employees who are available in either the general population or the percentage of protected class employees in a specific geographic area (Castetter & Young, 2000).
General population data encompass all individuals in a defined population, regardless of the qualifications that individuals possess. General labor market data encompass all individuals in a given labor market, whether qualified or unqualified for a position. These data provide critical information when determining whether discrimination has occurred or has failed to occur.

A variant on stock statistical analysis is referred to as concentration statistics, which are used to establish also a prima facia case for discrimination. Concentration statistics are used to illustrate specifically that significant differences exist between and among protected class groups, as distributed throughout various levels of an organization or as compared with other organizations (Castetter & Young, 2000). For example, in *O’Brien v. Sky Chefs, Inc.* (1982), the plaintiffs established a prima facia case of disparate impact by showing a significant disparity between the representation of women in lower paying jobs and in higher paying jobs.

The earliest court case involving a claim of employment discrimination is *Myart v. Motorola* (1964). In this case, a black applicant, with previous work experience, but who scored too low on a five-minute intelligence test, was denied employment as a television analyzer. The hearing officer maintained that the employer used the test score of the applicant unfairly by comparing the test score of the applicant to the test scores of more “advantaged” groups and that the test score failed to take into account factors contributing to cultural deprivation. Although the decision of the hearing officer was later overturned by the Supreme Court of Illinois for lack of evidence, a clear precedent was established to hear similar cases (Arvey & Faley, 1992).
Another significant unfair test discrimination case is *Griggs v. Duke Power Company* (1971). In this case, the defendant, who denied employment to the plaintiff by using the passage of two standardized tests as a substitute for a high school diploma, was accused of violating Title VII of the Civil Rights Act of 1964. In ruling on this case, The Supreme Court found that the employer assumes the burden of showing that a given job requirement has a manifest relationship to the job under consideration and; that, because a disproportionately high number of protected class applicants were excluded, the selection procedure was unfair (*Griggs v. Duke Power Company*, 1971).

Because the employment interview is categorized as a "test" according to the *Uniform Guidelines* (1978), the interview is subject to the same scrutiny as traditional pencil-paper employment tests. Although very little litigation has involved the face to face interview, there are several vulnerable areas that the interview exhibits. These areas of vulnerability include: using the same sex or the same race interviewers without prior interview training, inappropriate behaviors displayed by interviewers, discouraging protected class applicants during pre-interviewing, and past history of discrimination litigation (Dipboye, 1992).

One landmark case involving the employment interview is *Equal Employment Opportunity Commission v. Spokane Concrete Production Inc.* (1982). Here, both the time length of the interview and the preference, admitted by the interviewer, for a particular kind of applicant were called into question. In this case, a woman was denied employment after a 15-minute interview that focused on concerns of the employer over childcare issues rather than on the experience and qualifications of the applicant. The
court ruled that sexual discrimination occurred because the claims of the defendant represented a pretext for discrimination.

Another landmark case involving the employment interview is Phillips vs. Amoco Oil Co. (1982). In this case, the employer was found to be within legal bounds by excluding a black applicant based on the report of the interviewer that the applicant lacked the ability to learn aspects of the job. The court maintained further that the plaintiff had failed to illustrate a discriminatory intent on behalf of the defendant.

Understanding the significant role that state and federal legislation maintains in the employment process is necessary when insuring that an organizational selection process is not violating the law. Organizational administrators and school administrators must be cognizant of how explicit laws govern the selection of all employees, including teachers. Now that a foundation has been established, including discussion of the organizational staffing model, the macroanalytic and the microanalytic research perspectives, along with the legal considerations of the organizational staffing process, an examination of how organizational selection interview research has influenced the teacher selection interview research is appropriate.

**Influences of Organizational Selection Research on Teacher Selection Research**

Well documented is that selection research operated initially from a macroanalytic perspective. Researchers, who subscribed to the macroanalytic perspective focused on
assessing the reliability and the validity of the selection interview by using composite scores, also referred to as employability indices, as the unit of analysis. Investigators failed generally; however, to establish either significantly strong relationships between employability indices and job performance or to establish evidence that causation that might explain why employability indices were found to be either valid or invalid.

As a result of disappointingly low validity coefficients for employability indices, researchers began to examine the selection interview in component parts, rather than as a single events. The work of Mayfield and Carlson (1966) created a shift from the macroanalytic approach to the microanalytic approach when these researchers hypothesized that variables, unrelated to job performance, affected the decision making of interviewers; thus, causing validity coefficients to be skewed erroneously. Subsequent research identified race, sex, age, and physical disability as variables that influence the decisions of interviewers.

Selection research conducted within the educational setting, like selection research within the organizational setting, began with a macroanalytic approach. Researchers using the macroanalytic approach to teacher selection sought ultimately to validate the selection interview by attempting to maximize the amount of variance in teaching performance that could be accounted for by a predictor variable. Similar to the findings of organizational research from the macroanalytic perspective, teacher selection research from the macroanalytic perspective revealed that only a small percentage of variance associated with predictors of teaching performance was shared with measures of observed teaching performance (Harris, 1989; Mayfield & Carlson, 1966).
The limitations, revealed by using macroanalytic research, led investigators toward a microanalytic approach to teacher selection. Researchers using the microanalytic approach to teacher selection, according to Young and Ryerson (1986), attempted to "identify those variables that are unrelated to teaching performance and yet influence the selection decisions made by school administrators" (pg. 26). In identifying variables unrelated to a criterion measure of teaching performance, educational researchers followed the lead of organizational researchers who employed typically one of three different research strategies to investigate decision making at different stages of the teacher selection process: (1) field studies, (2) role-playing studies, and (3) resume studies.

One type of research strategy is field studies, which examine specific sources of potential bias during in an actual employment setting where predetermined interviewers and interviewees have been selected to represent the sources of potential bias under consideration. Another type of research strategy is role playing, which involves participants playing the part of an interviewer and an interviewee in a simulated employment setting where the sources of potential bias are manipulated either by a script or by special selection of participants (Young & Voss, 1986). Yet another type of research strategy is resume studies, in which administrators evaluate hypothetical applicants by reacting to simulated resumes that manipulate potential sources of bias. Resume studies are useful particularly when investigating the screening stage of the selection process (Gorman et al., 1978).
Screening decisions, using applicant resumes, have been examined from many different interviewer perspectives including superintendents, high school principals, elementary principals, and graduate students. Results obtained from resume studies suggested that screening decisions by administrators are influenced by factors such as sex, age, focal position under consideration, quality of information provided, and interactions among these factors. These findings have been consistent for elementary principals, for secondary principals, and for superintendents regardless of the geographic location of the sample chosen (Young, 1985).

Similar to the organizational selection process, educational administrators utilize interviews also to assess candidates under employment consideration. Administrator decisions, based on interview data, have been examined also by researchers in an attempt to understand what factors influence the selection decisions of administrators. The selection decisions of administrators, based on interview assessed data, have been found to be influenced by the information format used (Bolton, 1969), the structure of the interview (Young, 1983), and the interpersonal performance style of the applicants (Young, 1984).

As knowledge of the decision making process increased, within a teacher selection context, a framework for understanding and for assimilating past, present, and future research remained absent. Schalock (1979) stated, “As critical as teacher selection is to the field of education, no framework has emerged within which to synthesize the research that has been done or to guide future research” (p.369). In response to Schalock’s call for an assimilating “framework” for teacher selection research, Young
(1985) developed a structural model designed to synthesize existing teacher selection research and to guide further teacher selection research efforts.

The teacher selection model proposed by Young (1985) is referred to as the "shared variance paradigm" because the paradigm depicts variance associated with specific components of the teacher selection process. Figure 1 below represents the model developed by Young.

![Diagram of Young's shared variance model for teacher selection.]

Figure 1: Young's shared variance model for teacher selection.

Note: Set A represents variance associated with teaching performance. Set B represents variance associated with an interview decision. Set C represents variance associated with factors unrelated to teaching performance, but influential in decisions. Set D represents variance associated with screening decisions.
Young's model is illustrated by using a Venn diagram that depicts the selection process. Each of the four sets (A, B, C, and D) represents a specific component of the selection process. Set A represents variance associated with a criterion measure (teaching performance); set B represents variance associated with a predictor (interview decision); set C represents variance associated with factor(s) unrelated to job performance but influence interview decisions; and set D represents variance associated with a screening decision.

A visual representation of shared variance associated with teaching performance and interview decisions is depicted by $A \cap B$, and the amount of shared variance associated with interview decisions and factors unrelated to teaching performance that influence interview decisions is depicted by $B \cap C$. The intersection of set A and set B represents also the macroanalytic research perspective. Researchers sought to assess the relationship between set A (teaching performance) and set B (interview decisions) and sought to maximize the amount of variance associated with teaching performance which was accounted for by the variance associated with interview decisions ($A \cap B$).

From a microanalytic perspective, researchers focused on the amount of shared variance associated with interview decisions and factors unrelated to job performance that influenced interview decisions ($B \cap C$). Researchers sought to identify the influence of specific extraneous variables on employability indices using simulated resume and video studies. As a result, researchers identified variables unrelated to teaching
performance (set C) that influenced interview decisions and ultimately accounted for the previously identified variance in the interview decision (set B).

One very significant depiction in Young’s Shared Variance Paradigm is the lack of shared variance between set A and set C. Perhaps the most significant explanation for the lack of shared variance between set A and set C is that federal law forbids explicitly the interaction of these two variables. In other words, making interview decisions, based on protected class characteristics such as age, sex, race, or handicapping condition, are irrelevant to job performance and violate federal legislation.

Young’s model, in addition to synthesizing existing research and guiding future research, introduces another useful purpose for the Shared Variance Paradigm: the utility of the model. For example, Young (1983) treated the structure of the interview format (set C) as an extraneous variable in relation to the decisions made by interviewers (depicted in set B). Young found that “16% of the variance associated with decisions made in the selection interviews could accounted for by the interview format” (p.23).

As a result of Young’s finding, he proposed that factors unrelated to teaching performance be treated as suppressor variables. Young purported that the variance shared between a suppressor variable and the predictor variable (B∩C) be removed from the total amount of variance associated with the predictor. By removing B∩C (16% of the unwarranted variance) from the total variance, the utility of the selection interview is increased.

The Shared Variance Paradigm depicted graphic conceptualization of the teacher selection process by synthesizing existing teacher selection research and by guiding
future teacher selection research. Because decisions made by employers are continuous throughout the entire selection process, Young’s model is useful for understanding particularly what influences employment decisions during each phase of the teacher selection process. Two specific types of employer decisions are depicted graphically in Young’s model: (1) screening decisions and (2) interview decisions.

Screening decisions, discussed earlier from an organizational perspective, are made frequently during the teacher selection process. Each year, school administrators screen large numbers of teacher applicants, on the basis of paper credentials, in an attempt to create a smaller applicant pool. Young and Voss (1986) reported that over 99% of the administrators in the United States used the screening of paper credentials in reducing teacher applicant pools.

Among the various types of paper credentials submitted by applicants (e.g. application forms, letters of reference, and academic transcripts), resumes are a popular vehicle for assisting educational administrators in making employment decisions. Resume studies in the organizational literature revealed, for instance, that age bias existed in the screening of job applicants. Within an education context, several resume studies have been conducted examining the effects of variables such as age, sex, focal position, and work experience on the screening decisions of educational administrators.

One of the earliest teacher resume studies examining the effects of applicant age on employment decisions of administrators was conducted by Young and Allison (1982). In this study, the researchers sought to examine the effects of applicant age, applicant experience, and role of the administrator on the screening decisions of administrators.
The sample for this study was administrators in four mid-western states who were asked to examine paper credentials of teacher applicants. Results of the study indicated a main effect for the age of teacher candidates regardless of the candidate’s experience; younger candidates received higher ratings than did older candidates.

Young and Voss (1986), who raised concern regarding the generalizability of the Young and Allison (1982) study, used a national random sample to investigate administrator’s perceptions of teacher candidates. The three independent variables manipulated by Young and Voss (1986) were age (either 29 years old or 49 years old), teaching position (either chemistry or physical education), and quantity of resume information (either brief or complete) as represented on applicant resumes. The researchers discovered significant main effects for age and for focal position and discovered an interaction effect between age and focal position.

Further analysis of the interaction between age and focal position revealed that principals evaluated 49 year-old physical education teachers lower than 29 year-old chemistry teachers. According to Young and Voss (1986), “The interaction indicates that ratings that administrators assigned teacher candidates are a function of a specific age by focal position combination” (p. 39). Furthermore, a fundamental implication of this study is that results are more generalizable to a larger geographic area than reported by Young and Allison (1982).

In an attempt to advance further research on how applicant stimuli effect the employment decisions of school administrators, Young and McMurray (1986) investigated whether the source of applicant stimuli moderated the selection decisions
made by school administrators. Specifically, these researchers used resumes to manipulate applicant age (either 29 years old or 49 years old), focal position (either chemistry or physical education), quality of applicant information (undergraduate grade point average, either 3.50 or 2.68), and quantity of resume information (letter of recommendation, either brief or embellished). Young and McMurray were interested particularly in how the quality or the quantity of applicant stimuli moderated selection decisions.

Contrary to earlier research, Young and McMurray reported results that revealed a significant main effect for quality of information and revealed a significant interaction between age, position, and quality of information. Furthermore, the quality of applicant information was shown to have disparate impact for older candidates. Young and McMurray (1986) stated, “These data suggest the quality of candidate stimuli had a disparate impact for older candidates . . . Principals used different standards (GPA) for evaluating 49 year old chemistry candidates than for evaluating 29 year old candidates” (p.8).

Young and Joseph (1989) suspected that skill obsolescence was a confounding factor leading to the conflicting results obtained in previous resume studies. These investigators examined the effects of age, skill obsolescence (length of time since graduating college), quality of information, and focal position on the screening decisions made by principals. Young and Joseph sought to determine whether older candidates seeking chemistry teaching positions were perceived as technically competent as younger candidates seeking chemistry teaching positions.
Page 51 is missing
In yet another, more recent, study examining potential age bias during resume screening, Young et al. (1997) sought to determine, in two experiments, the impact of chronological age and perceived focal position demands of teacher candidates on the screening decisions of principals. The first experiment depicted teacher candidates as being either 29 years old or 49 years old, while the second experiment represented a control situation where no applicant age was provided. In both experimental conditions, different focal position configurations were used to operationalize perceived physical activity levels associated with particular focal positions.

Findings from the Young et.al. study indicated that older physical education teacher candidates were evaluated lower systematically than younger physical education teacher candidates on the criteria-specific ratings provided by principals. Chronological age of teacher candidates failed; however, to influence the assessments of principals when focal position configurations consisted of teacher positions other than physical education. Because this study failed to illustrate significant interaction between chronological age and focal positions other than physical education teacher, these researchers asserted that age discrimination might be less pronounced than existing research indicates.

Similar to the above-mentioned teacher selection resume studies, examining the impact of certain protected class characteristics on the screening decisions made by educational administrators, are teacher interview studies that investigated factors affecting the decisions made by educational administrators. To examine the impact certain variables have on the decision making of administrators during the interview
stage, investigators used role playing to simulate an interview setting. The role playing laboratory research, investigating employer decision making during an interview, was pioneered by Bolton (1969).

Bolton focused on information format, within the teacher interview context, in the following ways: (1) the instructional set given to interviewers, (2) the number of summary documents describing the teacher candidates, (3) the amount of information pertaining to the teacher candidates, and (4) the stimulus mode for receiving information about teacher candidates. This researcher discovered that interview decisions made by administrators were influenced by the informational format of the interview even though all teacher candidates were either equally qualified or equally unqualified for the position. His findings set the foundation for which subsequent interview research was conducted.

One subsequent research study, examining interview formats, was conducted by Young (1983), who manipulated two interview situations: (1) a dyad selection interview and (2) a panel selection interview. The dyad selection interview consisted of an administrator playing the part of the interviewer and a teacher playing the part of the teacher candidate, while the panel selection interview consisted of an administrator and a teacher role-playing interviewers and a teacher role-playing the teacher candidate. Young (1983) discovered that administrators were more likely to recommend the employment of teachers in dyad interviews than equally qualified candidates in panel interviews.
Another interview decision making study was conducted by Young (1984), who focused on the influence of interpersonal performance style of teacher candidates on selection decisions made by school administrators. The independent variables in this laboratory study consisted of three personal performance styles, as assessed for each interviewee after completing the Performance Style Test (PST). Administrators evaluated then all candidates using the candidate evaluation form, which contained three items: (1) overall interview performance, (2) job offer probability, and (3) decision certainty.

Young (1984) found that substantial variance existed in interview ratings which could be accounted for by interpersonal performance style of teachers. According to Young, “Teachers whose interpersonal performance could be characterized as adroit and Machiavellian tended to receive the highest interview ratings” (p.31). An important implication of this study is that administrators must be aware that interpersonal performance style influences their perceptions of the employability for teacher candidates.

Yet another study investigating the decision making of school administrators, within a teacher interview context, was conducted by Young and Pounder (1985). Given the consistent finding of age discrimination at the screening stage of teacher selection, Young and Pounder sought also to determine if age discrimination occurred during the interview stage of teacher selection and if this adverse impact could be moderated by the quality of applicant stimuli. These researchers conducted two separate experiments to
test their hypothesis: using (1) videotaped interviews and (2) simulated role-playing interviews.

To control for unwarranted variance due to position expectation, Young and Pounder (1985) provided all participants with job descriptions similar to the job descriptions used in earlier resume studies conducted by Young and Allison (1982) and by Young and Voss (1984). Furthermore, Young and Pounder used, as dependent variables, the same six evaluative resume criteria as employed by Young and Allison (1982). Results of the two experiments indicated that older candidates received equitable treatment during both the simulated role-playing interview experiment and the videotaped interview experiment.

Young and Pounder concluded that, while the resume criteria employed in the simulated interview study correlated strongly with administrators' perceptions of offering candidates employment, the employment perceptions of administrators failed to correlate with the chronological age of teacher candidates. Similarly, these researchers maintained that, while the resume criteria employed during the videotaped interview study correlated strongly with administrators' perceptions of offering candidates employment, the employment perceptions of administrators failed to correlate with the chronological age of teacher candidates. In sum, Young and Pounder raised questions as to the generalizability of candidate chronological age as a moderator for employment decisions of educational administrators.

Well established is that school administrators make continuous decisions as to which teacher candidates continue through the selection process and which teacher
candidates are removed from further selection consideration. Just as decisions are made about candidates during the screening stage of the selection process, decisions are made also during the interview stage of the selection process. One type of interview decision made by school administrators, during both the screening stage and the interview stage, is derived from the composite score of a structured interview such as the Teacher Perceiver Interview (TPI).

The Teacher Perceiver Interview: Historical Background and Advent

Shortly after the first review of selection interview research was completed (Wagner, 1949), Dr. Donald Clifton, a professor of educational psychology at the University of Nebraska, began investigating the behavioral characteristics of successful persons in the early 1950's (Selection Research Incorporated, p.5). Specifically, Clifton conducted a research project in which undergraduate seniors and graduate students served as counselors for freshman students, and Clifton asked freshman students to rate the experience with their respective counselor. This investigator noted that when a freshman student rated a counselor highly, other students assigned to the same counselor produced high ratings as well, while when a freshman student rated a counselor poorly, other students assigned to the same counselor produced poor ratings as well (Clifton & Hall, 1952).

As a result of Clifton's preliminary findings, subsequent research concluded that the counselor "as a person" was making the difference and that the personal values of the counselor differentiated a successful counselor from an unsuccessful counselor. In
assessing the different personal values of counselors, researchers noticed that counselors rated positively by students tended to focus their attention on individual students, while counselors rated negatively by students tended to focus more on process and on procedure (Selection Research Incorporated, p.6). As a result of how counselors viewed students, one researcher inferred that counselors rated positively had different thought patterns than the thought patterns of counselors rated negatively (Gaeddert, 1956).

In examining the thought patterns of counselors for both successful counselors and for unsuccessful counselors, researchers conducted interviews with all counselors, and a sorting process was used to identify the responses of successful counselors and of unsuccessful counselors. As thought patterns were identified and sorted, researchers began to develop an interviewing process for analyzing the thought patterns of individuals. The challenge for researchers was then identifying ways to apply this process to a practical situation.

In an attempt to discover practical applications for this interview process, researchers investigated whether the results of a teacher interview analysis would correlate with other evaluations of teachers. Bonneau (1956), for example, reported a .67 correlation between interviewer ratings of teachers and student ratings of the same teachers. Dodge and Clifton (1956) conducted a study also using the interview process for teachers and found significant correlation between interviewer ratings of teachers and subsequent student ratings of student teachers.

Another early study investigating the interview process as a predictor of teaching performance was conducted by Warner (1969), who reported,
"...a highly predictive relationship between the interview analysis at the conclusion of the teacher's senior year of college and the rating he/she received from administrators and students at the end of the teacher's first year of teaching" (as cited in Teacher Perceiver, 1997, p.4).

In examining a specific teaching focal position, Winesman (1969) reported that a highly predictive relationship existed between the results of a vocational teacher interview and the teacher-student rapport of vocational teachers. Although many of the above-mentioned studies are unpublished doctoral dissertations, some unretreivable, these reports were used to help support the belief that the performance of teachers during an interview was predictive of teaching performance.

Early studies examining the correlation between the performance of teachers during an interview and subsequent teaching performance led to the development of the Teacher Perceiver Interview (T.P.I.) in 1971. The company that developed the Teacher Perceiver Interview, Selection Research Incorporated (SRI) maintained that the Teacher Perceiver Interview is "an interview process for identifying the very best teaching talent. The process involves a structured, stress-free interview" yielding "results highly predictive of on-the-job-performance" (as cited in Miller et al., 1977, p.3).

The Teacher Perceiver Interview is a structured interview composed of sixty open-ended questions designed to provide applicants with an opportunity for self-expression with regard to various job-related issues. According to The Gallup Organization, the interview is then analyzed according to life themes that are defined as a "recurring and consistent pattern of thought, feeling, or behavior"(p.1). Each person
represents an aggregate of many different themes and this aggregate takes the form of a composite score, which is used to screen candidates.

The twelve themes used in the Teacher Perceiver Interview, grouped into three subsets, are as follows:

<table>
<thead>
<tr>
<th>Intrapersonal</th>
<th>Interpersonal</th>
<th>Extrapersonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission</td>
<td>Empathy</td>
<td>Individual</td>
</tr>
<tr>
<td>Investment</td>
<td>Rapport Drive</td>
<td>Input Drive</td>
</tr>
<tr>
<td>Focus</td>
<td>Listening</td>
<td>Activation</td>
</tr>
<tr>
<td>Perception</td>
<td>Objectivity</td>
<td>Gestalt</td>
</tr>
</tbody>
</table>

(see appendix A).

Each of the twelve “Teacher Themes” contains five open-ended questions that are asked to prospective or practicing teacher candidates. Responses provided by candidates are evaluated and scored, dichotomously, as either correct or incorrect using assigned values of either 0 or 1. Teacher candidates can, therefore, earn a score on the Teacher Perceiver Interview ranging from a low of 0 to a high of 60.
Since the development of the original 60-item interview in 1971, an abbreviated version of the Teacher Perceiver Interview emerged. Based on information gathered from a national study of the 60-item T.P.I., a 22-item “screener” was developed. This “screener,” formally known as the Teacher Perceiver Interview Screener (T.P.I.S.), measures ten of the twelve “Teacher Themes” mentioned in the above paragraph and consists of an item distribution as follows:

<table>
<thead>
<tr>
<th>Mission</th>
<th>(2 questions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathy</td>
<td>(3 questions)</td>
</tr>
<tr>
<td>Rapport Drive</td>
<td>(2 questions)</td>
</tr>
<tr>
<td>Individualized Perception</td>
<td>(2 questions)</td>
</tr>
<tr>
<td>Investment</td>
<td>(1 question)</td>
</tr>
<tr>
<td>Input Drive</td>
<td>(4 questions)</td>
</tr>
<tr>
<td>Activation</td>
<td>(3 questions)</td>
</tr>
<tr>
<td>Innovation</td>
<td>(2 questions)</td>
</tr>
<tr>
<td>Gestalt</td>
<td>(2 questions)</td>
</tr>
<tr>
<td>Objectivity</td>
<td>(1 question)</td>
</tr>
<tr>
<td>Total</td>
<td>(22 questions)</td>
</tr>
</tbody>
</table>

Although the Gallup Organization (1996) fails to assert a specific recommended “cut off” score for teacher applicants, the testing company does maintain that “A score of approximately 10 or more points on the 22-item teacher screener provides a conservative, yet efficient, reference point for further consideration of a candidate” (p.13). Still unclear is how the 22 items from the 60-item Teacher Perceiver Interview were selected as the basis for the T.P.I.S. Furthermore, a rationale as to why only ten of the twelve “Teacher Themes” were used in developing the T.P.I.S. has failed to emerge from either the research literature or from written request from the testing company.
Because both the T.P.I.S. and the T.P.I. involve an interviewer rating a teacher candidate on subjective criteria, significant interrater agreement must be achieved. Selection Research Incorporated stresses that all interviewers receive certification that over one hundred hours of study and instruction has been received and that a measure of 85 percent consistency has been achieved on item by item coding, when compared with Selection Research staff members. According to SRI, “Because the mission of education is so important and because the selection process itself is so complex, it is absolutely essential that those people using this technique be fully trained in the art and science of perceiving educational talent” (p.6).

The central purpose for training administrators as “perceivers” is to minimize or eliminate the subjective influences that sometimes interfere with the selection of candidates, a phenomenon SRI referred to as “glare.” The Teacher Perceiver selection process purports to strengthen objectivity in three ways: “(1) the interview is structured and every candidate is asked the same questions, in the same order (2) the interview promotes more spontaneous responses; which, in turn, minimizes the tendency for a candidate to respond “appropriately” and (3) responses of each candidate are judged against responses from outstanding teachers rather than interviewer biases” (SRI, p.7). Selection research Incorporated has maintained further that this structured process “in the hands of trained perceiver specialists minimizes subjective factors – mannerisms, taste, physical attractiveness or unattractiveness and the mood of the interviewer – which can interfere in the selection of educational personnel with the greatest success potential (p.17).
Validation Studies of the Teacher Perceiver Interview System

Shortly after the development of the first edition of the Teacher Perceiver Interview (1971), researchers sought to determine whether this newly created structured interview measured what it purported to measure. Millard and Brooks (1974) conducted a three-year study of the Teacher Perceiver in two school districts and found that

"of the 34 comparisons of ratings made by peers, administrators, and students, approximately 68 percent of them favor those groups who were most highly recommended by the Perceiver process while less than 15 percent favored those groups who were not as highly recommended" (as cited in Gallup, 1997, p.4).

Another study, conducted by Coker et al. (1976), unretrievable after exhaustive attempts, reported a significant difference between teachers rated high by administrators and teachers rated low by administrators and the total score on the Teacher Perceiver Interview (as cited in Miller, 1976).

Subsequent validation studies, produced also in the middle 1970's, reported significant correlation between the Teacher Perceiver Interview and student and administrator ratings of teachers. Also, several school districts in several areas of the country have conducted local validation studies and these studies indicate that the Teacher Perceiver Process correlates with other evaluations of teachers. Some studies, however, reported positive, but not significant results with regard to the Teacher Perceiver process (Simmons, 1976).
More recently, The Gallup Organization (1990) conducted a validity study of the Teacher Perceiver Interview process, unretrievable after several attempts, on a sample of 162 teachers from public and private school systems across the United States. The Teacher Perceiver Interview was administered to each teacher in the sample and each teacher was subsequently evaluated by school administrators. Each teacher was placed then, based on overall teaching proficiency, in one of four quartiles, as determined by each teacher's evaluating administrator.

The results of the Gallup (1990) study, as cited in promotional literature, indicated that teachers included in the top quartile had an average of 26.55 raw score points on the TPI; teachers placed in the second quartile had an average of 24.80 raw score points on the TPI; teachers placed in the third quartile had an average of 22.08 raw score points on the TPI, and those teacher perceived to be in the bottom quartile achieved an average raw score of 16.71 on the TPI. Researchers reported a correlation of .38, at the .05 level of probability, between high scores on the Teacher Perceiver Interview and high ratings of overall teaching performance by school administrators. Specifically, "teachers who received higher score on the TPI tended to be rated higher by administrators in terms of overall teaching proficiency" (as cited in The Gallup Organization, 1997, p.5).

Since 1990, The Gallup Organization has conducted two additional studies examining the validity of the Teacher Perceiver Interview. A 1994 Gallup Organization study used a sample of 211 teachers from, urban, suburban, and rural school districts and reported again a significant correlation between the TPI score of a teacher and the ranking received by the evaluating administrator. Another study correlated ratings of
over 20,000 students, in grades Four through Twelve, to quartile rankings made by principals, and these researchers found a .42 correlation.

One meta-analysis of the Perceiver Interview, unpublished and un retrievable, examined across a variety of organizational settings, was conducted by Schmidt (1993). This meta-analysis, which included purportedly only 12 teacher studies (n=380), revealed a correlation of .37 for the predictive validity of the Perceiver process. Schmidt, purporting that the validity of the Perceiver interview was generalizable across organizational settings stated,

"These findings show that the Perceiver Interview is a valid predictor of job performance ratings and sales success across a variety of jobs. The magnitude of the validity may vary with a particular type of job, but in the studies included in this meta-analysis the validity was fairly large and generalizable across situations, organizations, jobs, employers, etc." (Schmidt, 1993; as cited in Gallup Organization, 1997, p.6).

Unlike the full-version T.P.I., little research has been conducted the Teacher Perceiver Interview Screener. The Gallup Organization (1996), in an unpublished report, outlines the results of two national studies conducted in 1990 and in 1994. Gallup maintains that in the 1990 sample of 162 teachers, "the correlation to administrator’s quartile rating of teachers was .36" (p.13) and in the 1994 sample of 146 teachers the correlation was .39.

Another research study, conducted by Schmidt & Rader (1999), presented meta-analytic findings regarding interview validity. Included in this meta-analysis are four validation studies of teacher selection interviews that utilized dimensions of human behavior. Findings from this study suggest that a structured, twenty-two item interview
(not actually referred to as the Teacher Perceiver Interview Screener) yields validity coefficients ranging from .40-.51.

Now that an overview of the Teacher Perceiver Interview (predictor variable) has been provided, an examination of various criterion variables is needed to understand the relationship between the predictor variable and the criterion variable. When a selection decision is made to hire a classroom teacher, based on a Teacher Perceiver Interview performance score, a prediction is made regarding future teaching performance. To investigate the predictive validity of the Teacher Perceiver Interview, it becomes necessary first to consider the performance-based criterion measure used in explaining the extent to which a relationship exists between the predictor variable and the criterion variable.

The identification of an appropriate test for validation of performance criterion is critical because specific, mandated, federal employment procedures exist to ensure fairness when examining relationships between two employment variables. The Uniform Guidelines on Employee Selection Procedures (1978), outlined specific validation standards which included,

"Evidence of the validity of a test or other selection procedure by a criterion-related validity study should consist of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance (Section 1607.5, A)."

The specific types of criteria used for conducting validity studies are outlined also, in the "Technical Standards" section of the Uniform Guidelines, which reads,
"... Whatever criteria are used should represent important work behavior(s) or work outcomes. These criteria include but are not limited to production rate, error rate, tardiness, absenteeism, and length of service. A standardized rating of work performance may be used where a study of the job shows it is an appropriate criterion (Section 1607.14, B3).

Objective measures of job performance are grouped into two major categories: (1) production data and (2) personnel data. Production data includes measuring performance criteria such as the number of units produced or the dollar volume of sales. Personnel data is objective, equally to production data, and measures performance criteria such as absences, tenure, and rate of advancement.

Although objective measures of job performance are often easily obtainable, many job performance criterion measures are obtained through judgment ratings. These judgment ratings may consist of indirect ratings, such as grades issued to students by a classroom teacher, or may consist of direct ratings, such as the performance rating assigned to a classroom teacher by a school administrator. Direct ratings of job performance exist in three forms: (1) rating scales, (2) comparison systems, and (3) checklists.

In some instances, diagnostic rating scales are appropriate when assessing the relative strength or the relative weakness of a person’s job performance. Although rating scales measuring job performance exist in many forms, quantified rating scales with numerical anchors and with verbal anchors are preferred. Furthermore, scaled ratings using five anchor points has been considered common practice (Guion, 1956).
Rating scales do, however, exhibit some weaknesses, for which must be controlled. Systematic error variance, for instance, is often caused by response characteristics of the evaluator. Three common sources of rating error include (1) halo, (2) leniency, and (3) central tendency (Dipboye, 1992).

Perhaps the most common source of rating error is halo, which is the tendency for an evaluator to rate a person the same way on all traits because of an overall impression, whether favorable or unfavorable. Another source of rating error is leniency, which is the tendency for an evaluator to give ratings skewed systematically either favorable or unfavorable. Yet another source of rating error is central tendency, which is characterized by a restriction around the center of the rating scale (Arvey & Faley, 1992).

Sources of rating error, such as halo, leniency, and central tendency, are sometimes controlled for by using an employee comparison system. Comparison systems incorporate methods by which raters compare the overall effectiveness of one employee with another employee, or with the average of an employee group. Methods for using comparison systems include (1) rank order comparisons, (2) paired comparisons, (3) forced distribution comparisons, and (4) between-group comparisons.

Rank ordering is a process where a superior ranks subordinates by selecting first the best employee and selecting then the worst employee, and repeating the process with all remaining employees. Paired comparison involves ranking one employee with every other employee individually. The method of forced distribution involves an employer placing the names of employees in predetermined categories with corresponding proportions similar to a normal curve distribution. Finally, between-group employee
comparisons interpret an employee's rating in relation to the group in which the employee belongs.

According to Guion (1965), "Judgments make better criterion measure when they avoid glittering generalizations and get down to specific behavior (pg. 104). Evaluators will, therefore, utilize behavioral checklists to observe what an employee is doing, rather than actually assessing employee performance. These behavioral checklists exist in various forms, which include (1) summated ratings and (2) forced choice ratings.

Summated ratings refer to the total response weight assigned to verbally anchored critical incident categories. Forced choice ratings are part of a checklist system where evaluative statements are arranged in groups, from which the evaluator chooses the statements that are the most descriptive or the least descriptive of the individual employee being rated. In either case, rating has been shown to produce reliable and valid evaluations (Guion, 1965).

Employers are called on often to make assessments of employee performance based on observation and evaluation. These administrative ratings are most objective when there are no competing interests between the known purpose of the evaluation and the subsequent ratings provided by the evaluator. Furthermore, rater qualifications are tantamount in importance to the extent that the rater uses "first hand" information in judging performance and to the extent that an overall performance rating fails to represent one isolated observation.

Although criterion measures are not restricted to performance measures (Heneman et al., 1997), this study focuses on the predictive validity of teaching
performance and the predictors in this study are best assessed using performance criterion measures. Furthermore, teacher performance measures exist most often in the form of written observations and written evaluations conducted by school administrators (Gorton & Schneider, 1991). In assessing the relationship between a Teacher Perceiver Interview score and a subsequent teacher performance rating, careful consideration must be given to the methodology employed.
CHAPTER 3

MEDTHODOLOGY

To provide data addressing the research questions posed in Chapter One, two studies were conducted. Study I examines the Teacher Perceiver Interview Screener T.P.I.S.), and Study II examines the full-length, 60-question, Teacher Perceiver Interview (T.P.I.). The issue addressed specifically, in both studies, is the extent to which a predictive relationship exists between the Teacher Perceiver Interview score received by classroom teachers and subsequent performance measures of the same classroom teachers.
Study I

Subjects

The population for this investigation consists of 72 classroom teachers in a large school district in the southeastern United States. This school district serves approximately 9,400 students in grades K-12. District educational facilities include the following: 2 high schools, 3 middle schools, 11 elementary schools, and 1 alternative school.

This school system employs currently 650 certified staff members, 585 of whom are certified classroom teachers. Of the remaining 65 certified staff positions, 40 are building level administrative posts and 25 are central office administrative positions. The average teacher salary in this district for the 1999-2000 academic year is $38,420.00.

During the 1998 and 1999 school years, the Director of Personnel in this school system administered the Teacher Perceiver Interview Screener (T.P.I.S) to all teacher candidates deemed competitive after paper credentials had been reviewed. A total of 72 classroom teachers were hired subsequently during the 1998 and 1999 schools years collectively. With the exception of those teachers with complete T.P.I.S. profile sheets, no additional sampling criterion were chosen such as chronological age, ethnicity, teaching field, or gender; however, descriptive statistics on these data are provided in Chapter Four.
Predictor Variables

The predictor variables used in Study I are the individual scores and the composite scores of teachers derived from the 22-question Teacher Perceiver Interview Screener. Composite scores are calculated by summing the individual scores achieved for each of 10 Teacher Perceiver Themes. The distribution of questions within each teacher theme is as follows: (1) Mission: 2 questions, (2) Empathy: 3 questions, (4) Individualized Perception: 2 questions, (5) Investment: 1 question, (6) Input Drive: 4 questions, (7) Activation: 3 questions, (8) Innovation: 2 questions, (9) Gestalt: 2 questions, and (10) Objectivity: 1 question (see appendix A for definitions of each theme).

Subject responses to the 22 open-ended questions were evaluated by the Director of Personnel who coded “listen for” responses from subjects. “Listen for” responses are those responses, provided by interviewees, that contain specific words or phrases for which the “perceiver” is trained to recognize. Training in the coding of “listen for” responses is both required by and provided from the testing company.

To achieve status as a “perceiver,” the Director of Personnel received certification that required over 100 hours of study and instruction in the administration of the Teacher Perceiver Interview. This training focused on assisting “perceivers” in coding accurately “listen for” responses. Furthermore, the “perceiver” must achieve 85% consistency on item by item T.P.I. scoring when compared to Selection Research Incorporated (Gallup, Inc.) staff members (Schmidt & Rader, 1999).
After “listen for” responses are evaluated, the interviewer, or “perceiver,”
provides a numeric score for each of the 22 open ended questions. Responses provided
by candidates are scored dichotomously, as either correct or as incorrect using assigned
values of either 1 or 0. Teacher candidates can, therefore, earn a total T.P.I.S. score from
a low of 0 to a high of 22.

Criterion Variables

The two criterion variables chosen for this study are (1) ratings of the classroom
teacher by a school administrator and (2) the number of days the teacher missed work.
Ratings provided by principals were used to assess the extent to which the Teacher
Perceiver Interview Screener score earned by teachers was related to the subsequent job
performance of the same teachers. To control for the possible effects of rater bias, due to
such factors as self-serving bias (see Dipboye, 1992, p.36), the administrator completing
the teacher evaluation was in all cases different from the administrator conducting the
Teacher Perceiver Interview Screener.

As outlined by school district contract, and by state statute, every probational
classroom teacher is required to be observed at least two times annually by the immediate
supervisor of the teacher. Classroom teachers who maintain “probationary” status are
those teachers with four years or fewer of full-time classroom teaching experience. The
72 teachers that comprised the population for Study I are those teachers with one or fewer
years of experience.
Prescribed by school district contract also is the use of some form of performance appraisal instrument to be executed by the immediate supervisor of the classroom teacher. For elementary teachers, the immediate supervisor was, in all cases, the building principal. The immediate supervisor for middle school level teachers and for high school level teachers was, in all cases, either the assistant principal or the principal.

The existing teacher performance appraisal instrument used by immediate supervisors in this school system consists of a formative and a summative evaluation component that contained a total of eight performance standards. Performance criteria, varying in number, are described under each performance standard. Administrators would rate typically the presence or the absence of each performance standard, dichotomously, as either "yes" or "no."

To determine if the Teacher Perceiver Interview items could be aligned with items from the school district's existing teacher performance appraisal instrument, a pilot study was conducted. In this pilot study, teachers and administrators were asked to perform a Q-sorting activity (see Kerlinger, 1973) that required experienced classroom teachers and experienced school administrators to sort each of the 72 items from the teacher appraisal instrument into one, or more, or none, of the ten Teacher Perceiver Themes. Results of this sorting activity revealed zero agreement on sorting any of the 72 items into six of the 12 Teacher Perceiver Interview themes.

As a result of this finding, attention turned toward creating a rating form to be used by all evaluating principals, that would reflect more directly the T.P.I. Themes than the existing performance appraisal instrument. This newly created rating form consists of
10 questions, which merely re-state the prescribed T.P.I. theme in the form of a question (see appendix C). The rating scale for each T.P.I. theme was measured on a five-point, Likert-Type scale ranging from a “Low” of 1 to a “High” of 5.

Furthermore, principals were asked to use their extant knowledge as well as classroom observation data in rating the level of teacher performance for each Teacher Perceiver Theme. At no time was the supervising principal privy to the Teacher Perceiver Interview Screener score received by the teacher.

Another Criterion variable used in Study I was the number of days the teacher missed work since the date of hire. Each teacher in this study had 1 year of experience or less in the school system and absence was adjusted for time by calculating an average monthly absence. Only personal days and sick days were calculated for absenteeism.

Procedures – Actual Study

A letter of consent was secured from the Superintendent of this school system acknowledging participation in Study I. After securing the letter of consent from the Superintendent of Schools, exemption approval was sought and granted from The Human Subjects review Committee through The Ohio State University Research Foundation. This study was eligible for exemption from “formal” review because data needs for this study exist currently through the normal operating procedures of the school district.

After exemption from committee review, data retrieval began. Information retrieved from teacher personnel files, which are documents of public record, included (1) Teacher Perceiver Interview Screener themes scores and total interview scores (range=0-
22) for the participating teachers (n=72) and (2) the frequency of days absent from work since the date of hire. To protect employee anonymity, each teacher was identified by only a number (e.g., “teacher 1,” “teacher 2,” etc.) and only the interviewer, or “perceiver,” and the investigator were privy to the interview score of each teacher.

Immediate supervisors received then all necessary information and specific directions needed for completing the teacher rating form. First, supervising principals were given a definition for ten Teacher Perceiver Themes (see Appendix A). Second, supervising principals were asked to rate the level of performance exhibited by their respective evaluatees on the five-point scale described above.

Design and Analysis

The focus of Study I was examining the extent to which a predictive relationship existed between the Teacher Perceiver Interview Screener (T.P.I.S.) score of classroom teachers and the subsequent on-the-job performance evaluations of the same classroom teachers. The basic design for Study I was, therefore, a correlational design. Specifically, the degree of relationship between the variables in the research question was determined by using the Pearson Product-Moment Correlation Coefficient. Also employed in this study, along with the correlational analysis, was a regression analysis to develop a predictive model.
Study II

Subjects

The population for this investigation consists of 124 classroom teachers in a large suburban school district in the mid-western United States. This school district serves approximately 9,900 students in grades K–12. District educational facilities include: 2 high schools, 3 middle schools, and 10 elementary schools. The school system, one of the largest in the county, employs currently 771 certified staff members, 697 of who are certified classroom teachers. The average teacher salary in this district for the 1999-2000 academic year is $41,353.

Included also in the total number of certified staff members are 43 administrators. Central office administrators total 18, while building administrators number 25. Of the 25 building level administrators, 23 are building level principals and 2 are building level athletic directors.

During the 1997, 1998, and 1999 school years, the Coordinator of Human Resources for this school system administered the full-length, 60-question Teacher Perceiver Interview (T.P.I.) to all teacher candidates deemed competitive after paper credentials had been reviewed. A total of 139 classroom teachers were hired during the years 1997, 1998, and 1999 collectively. Because of missing interview data for 15 teachers, only 124 teachers could be used in this study. With the exception of those teachers with complete T.P.I. profile sheets, no additional sampling criterion were chosen such as chronological age, teaching field, ethnicity, or sex; however, descriptive statistics on these data are provided in Chapter Four.
Predictor Variables

The predictor variables used in Study II are the individual scores and the composite scores of teachers on the full-length, 60-question, Teacher Perceiver Interview (TPI). Summing all individual scores achieved for each of the 12 Teacher Perceiver Themes derived composite scores. The distribution of questions was arranged such that 5 open-ended questions, purporting to measure critical elements in achieving excellence in teaching, were asked for each of the following Teacher Perceiver Themes: (1) Mission, (2) Investment, (3) Focus, (4) Objectivity, (5) Listening, (6) Empathy, (7) Rapport Drive, (8) Activation, (9) Innovation, (10) Gestalt, (11) Input Drive, and (12) Individualized Perception (see appendix A for definitions of each theme).

Subject responses to the 60 open-ended questions were evaluated by the Coordinator of Human Resources who codes “listen for” responses from subjects. “Listen for” responses are those responses, provided by interviewees, that contain specific words or phrases for which the “perceiver” is trained to recognize. Training in the coding of “listen for” responses is required by and provided from the testing company.

To achieve status as a “perceiver,” the Coordinator of Human Resources received certification that required over 100 hours of study and instruction in the administration of the Teacher Perceiver Interview. This training is focused on assisting “perceiver” in coding accurately “listen for” responses. Furthermore, the “perceiver” must achieve 85%
consistency on item by item T.P.I. scoring when compared to Selection Research Incorporated (Gallup, Inc.) staff members (Schmidt & Rader, 1999).

After listen for responses are evaluated, the interviewer, or “perceiver,” provides a numeric score for each of the 60 open-ended questions. Responses provided by candidates are scored, dichotomously, as either correct or as incorrect using assigned values of either 0 or 1. Teacher candidates can, therefore, earn a total T.P.I. score from a low of 0 to a high of 60.

Criterion Variables

As in Study I, the two criterion variables in Study II are (1) ratings of the classroom by a school administrator and (2) the number of days the teacher missed work. Ratings provided by principals were used to assess the extent to which the Teacher Perceiver Interview Score earned by teachers was related to the subsequent job performance of the same classroom teachers. To control for the possible effects of rater bias, due to such factors as self-serving bias (see Dipboye, 1992, p. 36), the administrator completing the teacher observation was in all cases different from the administrator conducting the Teacher Perceiver Interview.

Similarly to Experiment I, and as outlined by school district contract, every probational classroom teacher is required to be observed at least two times annually by the immediate supervisor of the teacher. Classroom teachers who maintain “probationary” status are those teachers with three or fewer years of full-time classroom
teaching experience. For the purpose of this study, the population consisted of teachers with three or fewer years of teaching experience in the school district.

Prescribed by school district contract also is the use of some form of performance appraisal instrument to be executed by the immediate supervisor of the classroom teacher. For elementary teachers, the immediate supervisor was, in all cases, the building principal. The immediate supervisor for middle school level teachers and for high school level teachers was, in all cases, either the assistant principal or the principal.

The existing teacher performance appraisal instrument used by immediate supervisors in this school system consisted of a formative and a summative evaluation component that contained a total of six performance standards. Performance criteria, varying in number, were described under each performance standard. Administrators would rate typically the presence or the absence of each performance standard on a four-point, verbally anchored, rating scale labeled “unsatisfactory,” “needs improvement,” “does not apply,” and “satisfactory.”

The rating form used in Study I, to measure more directly the extent to which T.P.I. Themes were exhibited by classroom teachers, was used also in Study II. This rating form consists of 12 questions which merely state each T.P.I. theme in the form of a question (see appendix D). The rating scale for each T.P.I. theme was measured on a five-point Lickert-Type scale ranging from a “Low” of 1 to a “High” of 5.

Furthermore, principals were asked to use their extant knowledge as well as classroom observation data in rating the level of each Teacher Perceiver Theme. At no
time was the supervising principal privy to the Teacher Perceiver Score received by the teacher.

Another criterion variable used in Study II was the number of days the teacher reported absent from work since the start of the school year. In this study, the teachers had three or fewer years of experience in the school system and absence was adjusted for time by calculating a monthly average. As in Study I, only personal days and sick days were calculated in Study II.

Procedures – Actual Study

Similarly to Study I, Study II required securing a letter of consent from the Superintendent of this school system acknowledging participation in this study. After securing the letter of consent from the Superintendent of schools, exemption approval was sought and granted from The Human Subjects Review Committee through The Ohio State University Research Foundation. The Human Subjects Review Committee extended the initial exemption granted in Study I, to include Study II, because the protocol for both studies is the same.

After exemption from committee review, data retrieval began. Information retrieved from teacher personnel files, which are documents of public record, included (1) Teacher Perceiver Interview individual theme scores and total scores (range=0-60) for the participating teachers (n=124) and (2) the number of days the teacher missed work since the start of the school year. To protect employee anonymity, each teacher was identified
by only a number (e.g., “teacher 1,” “teacher 2,” etc.) and only the interviewer, or “perceiver,” and the investigator were privy to the interview score of each teacher.

Immediate supervisors received then all necessary information and specific directions needed for completing the teacher rating form. First, supervising principals were given a definition for each of the twelve Teacher Themes (see Appendix A). Second, supervising principals were asked to rank the level of performance exhibited by their respective evaluatees on the five-point scale described above.

Design and Analysis

The focus of Study II was examining the extent to which a predictive relationship existed between the full-length version, 60-question, Teacher Perceiver Interview (T.P.I.) score of classroom teachers and the subsequent on-the-job performance evaluations of the same classroom teachers. The basic design for Study II was, therefore, a correlational design. Specifically, the degree of relationship between the variables in the research question was determined by using the Pearson Product-Moment Correlation Coefficient.
CHAPTER 4

RESULTS

Chapter I presents an introductory description of the problem leading to this research project and the research questions examined within the scope of this study. Chapter II outlines related research on teacher selection, including the development of the Teacher Perceiver Interview (T.P.I) and the Teacher Perceiver Interview Screener (T.P.I.S). Chapter III describes the research procedures used in Study I and in Study II and includes details on research design and methodology. The present chapter summarizes the findings both from Study I and from Study II, and the results are presented as follow: (1) descriptive data of the target school district and of participants in each study, (2) data for the predictor variables, including item analyses for each T.P.I. theme, and (3) data illustrating relationships between the predictor variables and criteria variables.
Study I

The target population for Study I is teachers employed during the academic years of 1998 and 1999 by a large county school district located in southeastern state. At the time of data collection, this school system had an enrollment of approximately 9,400 students. Racial and ethnic data show that the largest percentage of these students (60.8%) is “Caucasian,” while 35.6% of the student enrollment is “Black,” and 3.6% of these students are “Indian, Asian, Hispanic, and Other.” Furthermore, these students attend school at an average annual rate of 94.1 percent.

This school system employs a total 1,285 employees, 681 of who are certified staff members. The total number of certified classroom teachers is 611. Per pupil expenditure in this district is $6,210, approximately 75% of which represents salaries and benefits of district employees. The average annual teacher salary in this district for the 1999-2000 academic year is $34,110. Table 1 indicates descriptive data for the school district used in Study I.
<table>
<thead>
<tr>
<th>Annual Average Daily Membership (ADM)</th>
<th>9,334</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expenditure Per Pupil</td>
<td>$6,210</td>
</tr>
<tr>
<td>Average Teacher Salary</td>
<td>$34,110</td>
</tr>
<tr>
<td>Student Attendance Rate</td>
<td>94.1 %</td>
</tr>
<tr>
<td>Staff Attendance Rate</td>
<td>96.8 %</td>
</tr>
</tbody>
</table>

**Student Racial Ethnic Percentages**

<table>
<thead>
<tr>
<th>Caucasian</th>
<th>Black</th>
<th>Indian, Hispanic, Asian, and Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>60.8 %</td>
<td>35.6 %</td>
<td>3.6 %</td>
</tr>
</tbody>
</table>

**Expenditures as a Percentage of Total Operating Expenses**

**Salary**

- Certified Staff: 48%
- Classified Staff: 12%

**Fringe Benefits**

- Certified: 12%
- Classified: 3%

**Combined Salary and Fringe Benefits (all): 75%**

---

*Table1: Descriptive data of school district for Study1.*
The target school district chosen for Study I employs a total of 681 certified employees, 611 of whom are certified classroom teachers. Of interest for Study I was 72 classroom hired during the 1998 and the 1999 school years.

Of these 72 classroom teachers, 64 are female and 8 are male. The average age of the participants is 31.4 years and the average number of years of teaching total teaching experience is 2.4 years. The “racial and ethnic” categorization of the participants is 62 “Caucasian” and 10 “other.” Demographic characteristics of the classroom teachers used in Study I are contained in Table 2.
<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td>72</td>
<td>31.41</td>
<td>8.12</td>
<td>22-54</td>
</tr>
<tr>
<td>Total Teaching Experience (Years)</td>
<td>72</td>
<td>2.40</td>
<td>3.96</td>
<td>0-23</td>
</tr>
<tr>
<td>District Teaching Experience (Years)</td>
<td>72</td>
<td>0.39</td>
<td>0.50</td>
<td>0-1</td>
</tr>
</tbody>
</table>

**Educational Level Attained (Frequencies)**

- Bachelor’s Degree: 58
- Master’s Degree: 13
- Master’s Degree + hours: 1

N=72

**Grade Level Teaching Assignments (Frequencies)**

- Elementary School: 40
- Middle School: 19
- High School: 13

- Sex: 64 Females
- 8 Males

- Race: 62 Caucasian
- 10 Other

Table 2: Demographic characteristics of classroom teachers in Study I.
Study II

The target population for this study was found in a large suburban school district in the mid-western United States. This school system has a district enrollment of approximately 9,900 students. Racial and ethnic data show that the largest percentage of these students (88.2%) is White, while students categorized as Asian and Multi-racial each represent 1.0% of the entire district enrollment. Furthermore, these students attend school at an average annual rate of 96.1 percent, which is 4.5 percent higher than the state average (EMIS, 1998).

This school system employs a total of 1,167 employees, 771 of whom are certified staff members. Per pupil expenditure in the district is $7,834, approximately 78 percent of which represents salaries for certified staff members. The average annual teacher salary in the district is $41,353.00 (EMIS, 1998). Descriptive data for the school system used in Study II are contained in Table 3.
Annual Average Daily Membership (ADM) 9,912
Expenditure Per Pupil $7,834
Average Teacher Salary $41,353
Student Attendance Rate 96.1%
Staff Attendance Rate 96.6%

Student Racial Ethnic Percentages

<table>
<thead>
<tr>
<th>White</th>
<th>Black</th>
<th>Hispanic</th>
<th>Asian</th>
<th>Multi-Racial</th>
</tr>
</thead>
<tbody>
<tr>
<td>88.2</td>
<td>2.2</td>
<td>1.0</td>
<td>7.6</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Expenditures as a Percentage of Total Operating Expenses

Salary

Certified Staff 50
Classified Staff 14

Fringes

Certified 10
Classified 4

Combined Salary and Fringes (all) 78

Table 3: Descriptive data of school district for Study II.
The target school district employs a total of 1,167 employees, 771 of who are certified staff members. Of those who are certified staff members, 697 represent classroom teachers. It was the total number of classroom teachers employed from 1997 to 1999 (N=139) that was targeted for this study. Of specific interest to the investigator, however, were those classroom teachers who had, as a part of their personnel file, a Teacher Perceiver Interview Profile Sheet. The Teacher Perceiver Profile Sheet delineates T.P.I. scores for each teacher according to each of the twelve T.P.I theme dimensions, as well as a composite interview score.

An examination of the personnel records indicated that T.P.I. profile sheets existed for 139 classroom teachers, so attention then focused on this target group. Unfortunately, 15 of the T.P.I profile sheets from the target group contained missing or incomplete data. The resulting and final population used for this study is, therefore, 124 classroom teachers.

Of these classroom teachers (N=124), 100 are female and 24 are male. The average age of the participants is 33.9 years and the average length of teaching experience in this school system is 2.4 years. The racial and ethnic categorization of the participants is 121 "white" and 3 "other." Table 4 contains demographic characteristics of the classroom teachers in Study II.
<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td>124</td>
<td>33.89</td>
<td>8.56</td>
<td>23-60</td>
</tr>
<tr>
<td>Total Teaching Experience (Years)</td>
<td>124</td>
<td>4.32</td>
<td>3.50</td>
<td>0-13</td>
</tr>
<tr>
<td>District Teaching Experience (Years)</td>
<td>124</td>
<td>2.44</td>
<td>1.05</td>
<td>0-3</td>
</tr>
</tbody>
</table>

**Educational Level Attained (Frequencies)**

- Bachelor’s Degree: 34
- Bachelor’s Degree + hours: 46
- Master’s Degree: 31
- Master’s Degree + hours: 11
- Ph.D.: 2

N=124

**Grade Level Teaching Assignments (Frequencies)**

- Elementary School: 48
- Middle School: 31
- High School: 45

**Sex**

- 100 Females
- 24 Males

**Race**

- 121 White
- 3 Other

Table 4: Demographic characteristics of classroom teachers for Study II.
Descriptive Statistics

The predictor variable used in Study I is the scores earned by teacher candidates on the abbreviated version of the Teacher Perceiver Interview, also referred to as the Teacher Perceiver Interview Screener (T.P.I.S.). Contained in Table 5 are descriptive statistics, reliabilities, and correlations for each predictor theme measured by the T.P.I.S. Descriptive statistics indicate the mean score earned by teacher candidates (N=72) on the 22-item T.P.I.S. as well as the respective standard deviation.

Reliability coefficients, as assessed by Cronbach Alpha, reveal the internal consistency of items comprising themes. These reliability coefficients range from a low of .00 (Empathy, Rapport, and Activation) to a high of .40 (Input Drive).
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Mission (2)(^a)</td>
<td>.97</td>
<td>.67</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II. Empathy (3)</td>
<td>.49</td>
<td>.56</td>
<td>-.04</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Rapport (2)</td>
<td>.96</td>
<td>.64</td>
<td>.05</td>
<td>.08</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV. Individualized</td>
<td>1.01</td>
<td>.71</td>
<td>-.20</td>
<td>-.03</td>
<td>-.10</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI. Investment (1)</td>
<td>.38</td>
<td>.49</td>
<td>-.01</td>
<td>.15</td>
<td>-.02</td>
<td>.10</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII. Input Drive (4)</td>
<td>3.07</td>
<td>.98</td>
<td>-.09</td>
<td>.25</td>
<td>-.05</td>
<td>-.11</td>
<td>.21</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIII. Activation (3)</td>
<td>2.11</td>
<td>.78</td>
<td>.03</td>
<td>.00</td>
<td>.01</td>
<td>-.04</td>
<td>-.07</td>
<td>-.04</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX. Innovation (2)</td>
<td>1.51</td>
<td>.65</td>
<td>.00</td>
<td>-.08</td>
<td>.04</td>
<td>-.06</td>
<td>.01</td>
<td>.26</td>
<td>.14</td>
<td>.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. Gestalt (2)</td>
<td>1.35</td>
<td>.44</td>
<td>-.11</td>
<td>.22</td>
<td>.07</td>
<td>.10</td>
<td>.11</td>
<td>-.06</td>
<td>.03</td>
<td>.01</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>XI. Objectivity (1)</td>
<td>.38</td>
<td>.49</td>
<td>.12</td>
<td>.05</td>
<td>.03</td>
<td>-.14</td>
<td>-.01</td>
<td>.24</td>
<td>.19</td>
<td>.32</td>
<td>-.03</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 5: Descriptive statistics, reliabilities, and correlations for each predicted T.P.I. theme.

\(^a\)Number in parentheses indicates the number of items contained within each T.P.I. theme.

\(^b\)Coefficients on the diagonal are reliabilities assessed by Cronbach Alpha.

Note: \(n=72\)

To assess the integrity of the items comprising each T.P.I. theme, part whole correlations (McNemar, 1968) were computed. These part-whole correlations illustrate what each theme item contributes to its respective theme independent of that item’s contribution to the composite score. Displayed in Tables 6-13 are part-whole correlations and correlations between the items comprising each T.P.I. theme and all other themes.
<table>
<thead>
<tr>
<th></th>
<th>I&lt;sup&gt;b&lt;/sup&gt;</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.02&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-.07</td>
<td>.02</td>
<td>-.15</td>
<td>.12</td>
<td>-.09</td>
<td>.02</td>
<td>.11</td>
<td>-.08</td>
<td>.24</td>
</tr>
<tr>
<td>13</td>
<td>.02</td>
<td>.02</td>
<td>.04</td>
<td>-.14</td>
<td>-.14</td>
<td>-.05</td>
<td>.02</td>
<td>-.11</td>
<td>-.08</td>
<td>-.08</td>
</tr>
</tbody>
</table>

Table 6: Analysis of items comprising Mission.

<sup>a</sup>Numbers in italics represent individual items comprising a T.P.I.S. theme.
<sup>b</sup>Roman numerals indicate T.P.I.S. themes: I. Mission, II. Empathy, III. Rapport Drive, IV. Individualized Perception, VI. Investment, VII. Input Drive, VIII. Activation, IX. Innovation, X. Gestalt, XI. Objectivity.
<sup>c</sup>Coefficients in bold represent T.P.I.S. theme part-whole correlations.

Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

An examination of Table 6 reveals relatively low correlations among items (1, 13) and themes. When the composite score for Mission is corrected for an individual item’s contribution, through a part-whole correlational technique, neither item 1 nor item 13 contributes meaningful variance to the assigned theme ($r^2 = .04$ and $r^2 = .04$, respectively). In fact, item 1 contributes substantially more variance to Objectivity ($r^2 = 5.8\%$) than item 1 contributes to Mission. Likewise, the variance shared between item 13 (corrected through part-whole correlation) is less than the variance shared between item 13 and themes defined as Individualized Perception and Investment.
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>-.14</td>
<td>-.04</td>
<td>.03</td>
<td>-.03</td>
<td>-.12</td>
<td>.20</td>
<td>-.05</td>
<td>-.17</td>
<td>.32</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>.22</td>
<td>.09</td>
<td>.12</td>
<td>.08</td>
<td>.13</td>
<td>.06</td>
<td>-.03</td>
<td>.05</td>
<td>.07</td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>.01</td>
<td>-.17</td>
<td>.01</td>
<td>-.12</td>
<td>-.02</td>
<td>.13</td>
<td>.15</td>
<td>.16</td>
<td>-.25</td>
<td>.13</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Analysis of items comprising Empathy.

*Numbers in italics represent individual items comprising a T.P.I.S. theme.
Coefficients in bold represent T.P.I.S. theme part-whole correlations.
Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

For the items purporting to measure Empathy (26, 38, & 50), only item 50 shares a slight amount of variance ($r^2=2.9\%$) with its designated theme (Empathy). Without exception, all items purporting to comprise the Empathy theme were found to share more variance with an undesignated theme than with the designated theme. To illustrate, item 26 shares more variance with Gestalt ($r^2=10.2\%$) than Empathy ($r^2=0.2\%$), item 38 shares more variance with Mission ($r^2=4.8\%$) than with Empathy ($r^2=0.8\%$), and item 50 shares more variance with Gestalt ($r^2=6.3\%$) than with Empathy ($r^2=2.9\%$).
Table 8: Analysis of items comprising Rapport Drive.

Numbers in italics represent individual items comprising a T.P.I.S. theme.
Coefficients in bold represent T.P.I.S. theme part-whole correlations.
Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

As illustrated in Table 8, items purporting to comprise Rapport Drive (3, 39) contribute the same amount of variance with the target theme (0.8 %). Furthermore, both of these themes contribute more variance to Empathy than these items contribute to Rapport Drive. Specifically, item 3 contributes substantially more variance to Empathy (6.8%) than item 3 contributes to Rapport Drive and item 39 contributes more variance to Empathy (2.9%) than to the target theme. Observable also is that item 3 contributes more variance to five themes, in addition to Empathy (Mission, Individualized Perception, Innovation, Gestalt, and Objectivity) than item 3 contributes to Rapport Drive.
Table 9: Analysis of items comprising Individualized Perception.

*Numbers in italics represent individual items comprising a T.P.I.S. theme.
*Coefficients in bold represent T.P.I.S. theme part-whole correlations.
Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

Revealed in Table 9 is that the two items purporting to measure Individualized Perception (28, 4) fail to share meaningful variance (0.01%) with the target theme. In fact, items comprising Individualized Perception contribute more variance to all themes other than Individualized Perception than these items contribute to the target theme. For example, item 28 contributes substantially more variance to Objectivity (7.8%) than item 28 contributes to Individualized Perception. Observed also is that the variance shared between item 4 and Individualized Perception (0.01%) is less than the variance shared between item 4 and Rapport Drive (4.0%). A similar pattern is found, with smaller amounts of shared variance, for all relationships among items and themes.
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>-.21</td>
<td>.26</td>
<td>-.12</td>
<td>.03</td>
<td>.21</td>
<td>.30</td>
<td>-.11</td>
<td>.04</td>
<td>.01</td>
<td>.21</td>
</tr>
<tr>
<td>31</td>
<td>.15</td>
<td>.20</td>
<td>-.02</td>
<td>-.20</td>
<td>.22</td>
<td>.23</td>
<td>.07</td>
<td>.15</td>
<td>-.11</td>
<td>.15</td>
</tr>
<tr>
<td>43</td>
<td>-.12</td>
<td>-.11</td>
<td>.07</td>
<td>-.06</td>
<td>-.03</td>
<td>.13</td>
<td>-.06</td>
<td>.09</td>
<td>.02</td>
<td>.04</td>
</tr>
<tr>
<td>55</td>
<td>-.07</td>
<td>.25</td>
<td>-.07</td>
<td>-.03</td>
<td>.12</td>
<td>.27</td>
<td>-.01</td>
<td>.27</td>
<td>-.05</td>
<td>.18</td>
</tr>
</tbody>
</table>

Table 10: Analysis of items comprising Input Drive.

*Numbers in italics represent individual items comprising a T.P.I.S. theme.
Coefficients in bold represent T.P.I.S. theme part-whole correlations.
Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

An examination of Table 10 reveals that when the composite score for Input Drive is corrected for an individual item’s contribution, through a part-whole technique, all items comprising Input Drive contribute substantially more variance to the target theme than to any other theme. Specifically, items 7, 31, 43, and 55 contribute meaningful variance to Input Drive (9.0%, 5.3%, 1.7%, and 7.3 % respectively).
Table 11: Analysis of items comprising Activation.

*Numbers in italics represent individual items comprising a T.P.I.S. theme.
Coefficients in bold represent T.P.I.S. theme part-whole correlations.
Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

For the items purporting to measure Activation (20, 32, and 56), all items fail to contribute meaningful variance (0.01, 0.40, 0.09, respectively) to the target theme (Activation). Without exception, all items comprising the Activation theme were found to share more variance with undesignated themes than with the designated theme. For illustration, item 20 shares more variance with Objectivity (4.0%) than with Activation (0.01%), item 32 shares more variance with Input Drive (1.7%) than with Activation (0.40%), and item 56 shares more variance with Objectivity (3.2%) than with Activation (0.09%).
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>-.02</td>
<td>.09</td>
<td>.03</td>
<td>-.12</td>
<td>-.01</td>
<td>.27</td>
<td>.02</td>
<td>.16</td>
<td>.00</td>
<td>.29</td>
</tr>
<tr>
<td>9</td>
<td>.02</td>
<td>-.18</td>
<td>.03</td>
<td>.01</td>
<td>.02</td>
<td>.12</td>
<td>.17</td>
<td>.16</td>
<td>.01</td>
<td>.20</td>
</tr>
</tbody>
</table>

Table 12: Analysis of items comprising Innovation.

<sup>a</sup>Numbers in italics represent individual items comprising a T.P.I.S. theme.
<sup>b</sup>Roman numerals indicate T.P.I.S. themes: I. Mission, II. Empathy, III. Rapport Drive, IV. Individualized Perception, VI. Investment, VII. Input Drive, VIII. Activation, IX. Innovation, X. Gestalt, XI. Objectivity.
<sup>c</sup>Coefficients in bold represent T.P.I.S. theme part-whole correlations.

Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

As illustrated in Table 12, items comprising Innovation (45, 9) share a slight amount of variance with the target theme (2.6%). When the composite score for Innovation is corrected, through a part-whole correlational technique, both items comprising Innovation (45, 9) contribute more variance to Objectivity (8.4% and 4.0% respectively) than to Innovation (2.6%). Furthermore, item 9 shares more variance with Empathy (3.2%) and with Activation (2.9%) than with Innovation (2.6%).
Table 13: Analysis of items comprising Gestalt.

*Numbers in italics represent individual items comprising a T.P.I.S. theme.
*Coefficients in bold represent T.P.I.S. theme part-whole correlations.
Note: T.P.I.S. themes V. Listening and XII. Focus are not themes measured by the T.P.I.S.

An examination of Table 13 reveals relatively low part-whole correlations among Gestalt items (58, 22) and themes. Item 22, however, represents an outlier in that this item contributes substantially more variance (5.3%) to Empathy than with any other theme. In general, items 58 and 22 contribute a very small proportion of variance (.50%) to Gestalt. Furthermore, item 58 contributes more variance to Empathy (0.8%), Individualized Perception (0.8%), and Investment (0.6%), than item 58 contributes to Gestalt. Notice also that item 22 contributes more variance to Empathy (5.3%) than item 22 contributes to Gestalt.
The two criterion measures selected for Study I are (1) ratings provided by each teacher’s evaluating principal and (2) the number of days the teacher missed work. To assess the relationship between interview scores and the above-mentioned criterion variables, a correlational analysis was conducted. As illustrated in Table 14, the highest degree of relationship between the predictor variable and rating variable is the 8.4 percent of variance shared between the principals’ rating of Activation performance and the predicted scores attained for the Activation theme. Revealed also in Table 14 is that the highest degree of relationship between a predictor variable and the absenteeism variable is the 5.3% of variance shared between absenteeism and predicted scores attained for the Rapport theme.
<table>
<thead>
<tr>
<th></th>
<th>R^a</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>Absenteeism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
<td>IV</td>
<td>V</td>
<td>VI</td>
<td>VII</td>
<td>VIII</td>
<td>IX</td>
<td>X</td>
<td>XI</td>
</tr>
<tr>
<td>I. Mission (2)b</td>
<td>.12</td>
<td>.21</td>
<td>.06</td>
<td>-.11</td>
<td>.07</td>
<td>-.10</td>
<td>.03</td>
<td>-.03</td>
<td>.01</td>
<td>.14</td>
<td>.02</td>
</tr>
<tr>
<td>II. Empathy (3)</td>
<td>-.12</td>
<td>-.04</td>
<td>-.09</td>
<td>-.16</td>
<td>-.13</td>
<td>-.26</td>
<td>-.30</td>
<td>-.21</td>
<td>-.07</td>
<td>-.03</td>
<td>.17</td>
</tr>
<tr>
<td>III. Rapport (2)</td>
<td>-.11</td>
<td>-.01</td>
<td>-.18</td>
<td>-.14</td>
<td>-.02</td>
<td>-.21</td>
<td>-.13</td>
<td>-.01</td>
<td>-.13</td>
<td>.04</td>
<td>.23</td>
</tr>
<tr>
<td>IV. Individualized Perception (2)</td>
<td>-.07</td>
<td>.04</td>
<td>.12</td>
<td>-.02</td>
<td>.00</td>
<td>.02</td>
<td>-.03</td>
<td>-.08</td>
<td>.02</td>
<td>.04</td>
<td>.05</td>
</tr>
<tr>
<td>V. Investment (1)</td>
<td>.21</td>
<td>.07</td>
<td>.05</td>
<td>.11</td>
<td>.09</td>
<td>.01</td>
<td>-.03</td>
<td>.02</td>
<td>.08</td>
<td>-.02</td>
<td>.21</td>
</tr>
<tr>
<td>VII. Input Drive (4)</td>
<td>.22</td>
<td>.03</td>
<td>.13</td>
<td>.04</td>
<td>.09</td>
<td>.10</td>
<td>.10</td>
<td>.11</td>
<td>.10</td>
<td>.06</td>
<td>-.08</td>
</tr>
<tr>
<td>VIII. Activation (3)</td>
<td>.26</td>
<td>.18</td>
<td>.21</td>
<td>.10</td>
<td>.17</td>
<td>.20</td>
<td>.29</td>
<td>.27</td>
<td>.22</td>
<td>.22</td>
<td>-.14</td>
</tr>
<tr>
<td>IX. Innovation (2)</td>
<td>.29</td>
<td>.11</td>
<td>.08</td>
<td>.07</td>
<td>.14</td>
<td>.20</td>
<td>.30</td>
<td>.25</td>
<td>.12</td>
<td>-.07</td>
<td>-.04</td>
</tr>
<tr>
<td>X. Gestalt (2)</td>
<td>.15</td>
<td>.19</td>
<td>.12</td>
<td>.04</td>
<td>.21</td>
<td>.05</td>
<td>.08</td>
<td>-.04</td>
<td>.24</td>
<td>.15</td>
<td>.02</td>
</tr>
<tr>
<td>XI. Objectivity (1)</td>
<td>.09</td>
<td>.04</td>
<td>.05</td>
<td>-.04</td>
<td>.05</td>
<td>.08</td>
<td>.16</td>
<td>.05</td>
<td>.01</td>
<td>-.02</td>
<td>.02</td>
</tr>
<tr>
<td>Absenteeism</td>
<td>-.29</td>
<td>-.05</td>
<td>-.17</td>
<td>-.04</td>
<td>-.19</td>
<td>-.18</td>
<td>-.26</td>
<td>-.18</td>
<td>-.18</td>
<td>-.17</td>
<td>-.</td>
</tr>
</tbody>
</table>

Table 14: Relationship between predictor and criterion measures.

^a Bolded column headings represent administrator ratings of teaching performance by T.P.I. theme.

^b Number in parentheses indicates the number of items contained within each T.P.I. theme.

Note: T.P.I. themes Listening and Focus are not measured by the T.P.I.S
To determine if a linear combination of predictor variables could account for a meaningful amount of variance associated with absenteeism, a regression analysis was conducted. This analysis utilizes a full model with all 10 themes regressed against the criterion variable absenteeism using a blockwise entry. Contained in Table 15 is the omnibus test of the regression model.

Collectively, these themes account for 16% of the variance associated with absenteeism. This amount of variance improves considerably the total variance accounted for relative to the individual contribution of a single variable. In fact, the largest amount of variance accounted for by a single variable is 4.80% (as shown in Table 16). The individual contributions of the separate themes, given the other themes in the model, are found in Table 16. The most important themes of this particular linear combination are Empathy and Rapport, while the least important themes for this linear combination are Mission and Innovation.

An examination of the direction of loadings (see Standardized Beta Coefficients in Table 16) reveals some interesting findings. High scores on Empathy and Rapport are more likely to be associated with high absenteeism rates. In a broad sweep, these findings may suggest that an individual’s concern for developing favorable relationships with others may be internalized to self. That is, individuals who have a need for favorable relationships with others may fulfill this need by missing work.

This interpretation is somewhat supported by the negative relationship between Input Drive and Absenteeism (see Standardized Beta Coefficients in Table 16). Teachers that seek to find materials and relevant resources to enhance student achievement (high
Input Drive) are more likely to attend work than those teachers with lower Input Drive scores.
<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>4.46</td>
<td>10</td>
<td>.45</td>
<td>1.18</td>
<td>.32</td>
</tr>
<tr>
<td>Residual</td>
<td>23.14</td>
<td>61</td>
<td>.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>27.61</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 15: Full Regression model containing all themes.

<table>
<thead>
<tr>
<th>Variables</th>
<th>b</th>
<th>Beta</th>
<th>T Test</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission</td>
<td>-2.16E-03</td>
<td>.00</td>
<td>-0.02</td>
<td>.18</td>
</tr>
<tr>
<td>Empathy</td>
<td>.20</td>
<td>.18</td>
<td>1.42</td>
<td>.99</td>
</tr>
<tr>
<td>Rapport</td>
<td>.22</td>
<td>.22</td>
<td>1.85</td>
<td>.16</td>
</tr>
<tr>
<td>Individualized Perception</td>
<td>4.01E-02</td>
<td>.05</td>
<td>.38</td>
<td>.07</td>
</tr>
<tr>
<td>Investment</td>
<td>.28</td>
<td>.22</td>
<td>1.75</td>
<td>.71</td>
</tr>
<tr>
<td>Input Drive</td>
<td>-.12</td>
<td>-.18</td>
<td>-1.35</td>
<td>.09</td>
</tr>
<tr>
<td>Activation</td>
<td>-.11</td>
<td>-.14</td>
<td>-1.18</td>
<td>.18</td>
</tr>
<tr>
<td>Innovation</td>
<td>1.065E-02</td>
<td>.01</td>
<td>.08</td>
<td>.24</td>
</tr>
<tr>
<td>Gestalt</td>
<td>-9.780E-02</td>
<td>-.07</td>
<td>-.56</td>
<td>.93</td>
</tr>
<tr>
<td>Objectivity</td>
<td>9.649E-02</td>
<td>.08</td>
<td>.58</td>
<td>.58</td>
</tr>
<tr>
<td>Constant</td>
<td>.65</td>
<td></td>
<td>1.36</td>
<td>.56</td>
</tr>
</tbody>
</table>

Table 16: Theme contribution to regression model.

Note: R = .40
R² = .16
The predictor variable used in Study II is the scores earned by teacher candidates on the complete, 60-question, Teacher Perceiver Interview (T.P.I.). Contained in Table 15 are descriptive statistics, reliabilities, and correlations for each predicted theme measured by the T.P.I. Descriptive statistics indicate the mean score achieved by teacher candidates (N=124) on the T.P.I. as well as the respective standard deviation. Reliability coefficients, as assessed by Cronbach Alpha, reveal the internal consistency of items comprising interview themes. These reliability coefficients range from a low of .00 (Input Drive) to a high of .66 (Investment).
<table>
<thead>
<tr>
<th>Theme</th>
<th>Mean</th>
<th>SD</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Mission (5)*</td>
<td>2.9</td>
<td>1.1</td>
<td>.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II. Investment (5)</td>
<td>3.9</td>
<td>1.1</td>
<td>-.05</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Focus (5)</td>
<td>3.8</td>
<td>1.1</td>
<td>.12</td>
<td>.27</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV. Objectivity (5)</td>
<td>2.8</td>
<td>1.2</td>
<td>.01</td>
<td>.16</td>
<td>.13</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V. Listening (5)</td>
<td>2.6</td>
<td>1.2</td>
<td>.06</td>
<td>-.06</td>
<td>.04</td>
<td>.18</td>
<td>.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI. Rapport (5)</td>
<td>2.8</td>
<td>1.2</td>
<td>.05</td>
<td>-.02</td>
<td>.17</td>
<td>.30</td>
<td>.26</td>
<td>.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII. Empathy (5)</td>
<td>2.4</td>
<td>1.2</td>
<td>.19</td>
<td>.01</td>
<td>.13</td>
<td>.19</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIII. Activation (5)</td>
<td>2.9</td>
<td>1.1</td>
<td>.03</td>
<td>.15</td>
<td>.04</td>
<td>.17</td>
<td>.24</td>
<td>.18</td>
<td>.25</td>
<td>.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX. Innovation (5)</td>
<td>3.1</td>
<td>1.0</td>
<td>-.13</td>
<td>.07</td>
<td>-.03</td>
<td>.01</td>
<td>.10</td>
<td>-.03</td>
<td>.18</td>
<td>.17</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X. Gestalt (5)</td>
<td>2.8</td>
<td>1.5</td>
<td>.09</td>
<td>.05</td>
<td>.06</td>
<td>.05</td>
<td>.02</td>
<td>.18</td>
<td>.01</td>
<td>-.04</td>
<td>.06</td>
<td>.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XI. Input Drive (5)</td>
<td>3.2</td>
<td>1.0</td>
<td>.22</td>
<td>-.13</td>
<td>.10</td>
<td>.11</td>
<td>.03</td>
<td>.01</td>
<td>.12</td>
<td>-.01</td>
<td>-.01</td>
<td>-.13</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>XII. Individualized Perception (5)</td>
<td>2.9</td>
<td>1.2</td>
<td>-.03</td>
<td>.11</td>
<td>.08</td>
<td>.22</td>
<td>.26</td>
<td>.14</td>
<td>.18</td>
<td>.17</td>
<td>.19</td>
<td>.04</td>
<td>.04</td>
<td>.20</td>
</tr>
</tbody>
</table>

Table 17: Descriptive statistics, reliabilities, and correlations for each predicted T.P.I theme

*Number in parentheses indicates the number of items contained within each T.P.I. Theme.

bCoefficients on the diagonal are reliabilities assessed by Cronbach Alpha.

Note: n = 124.
To assess the integrity of the items comprising each T.P.I. theme, part whole correlations (McNemar, 1968) were computed. These part-whole correlations illustrate what each theme item contributes to its respective theme independent of that item’s contribution to the composite score. Displayed in Tables 18-29 are part-whole correlations and correlations between the items comprising each T.P.I. theme and all other themes.
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>.13</td>
<td>.01</td>
<td>.08</td>
<td>-.05</td>
<td>.02</td>
<td>-.05</td>
<td>.18</td>
<td>.09</td>
<td>.05</td>
<td>-.08</td>
<td>-.08</td>
<td>-.07</td>
</tr>
<tr>
<td>13</td>
<td>.06</td>
<td>.13</td>
<td>.06</td>
<td>-.02</td>
<td>.09</td>
<td>-.14</td>
<td>.20</td>
<td>-.09</td>
<td>.00</td>
<td>-.21</td>
<td>.06</td>
<td>-.03</td>
</tr>
<tr>
<td>25</td>
<td>.00</td>
<td>.08</td>
<td>.02</td>
<td>-.15</td>
<td>-.18</td>
<td>.18</td>
<td>.06</td>
<td>-.04</td>
<td>-.18</td>
<td>.09</td>
<td>.10</td>
<td>.07</td>
</tr>
<tr>
<td>37</td>
<td>.48</td>
<td>.12</td>
<td>.10</td>
<td>.05</td>
<td>.17</td>
<td>-.02</td>
<td>-.07</td>
<td>.03</td>
<td>-.01</td>
<td>.01</td>
<td>.07</td>
<td>.25</td>
</tr>
<tr>
<td>49</td>
<td>.54</td>
<td>.09</td>
<td>-.13</td>
<td>.07</td>
<td>.02</td>
<td>-.03</td>
<td>.12</td>
<td>.07</td>
<td>.11</td>
<td>-.06</td>
<td>.08</td>
<td>.07</td>
</tr>
</tbody>
</table>

Table 18: Analysis of items comprising Mission.

*Numbers in italics represent individual items comprising a T.P.I. Theme.


*Coefficients in bold represent T.P.I. theme part-whole correlations

Considering the items purporting to measure Mission (1, 13, 25, 37, and 49) Table 16 reveals that items 37 and 49 share a moderate amount of variance ($r^2 = 23\%$, and $29\%$ respectively) with the target theme when the effects of item 25 is removed from the component. Item 25, however, fails to share any degree of relationship with the Mission theme it purports to measure. Furthermore, item 25 contributes more variance to all undesignated themes than item 25 contributes to the designated theme. Also, item 13 contributes more variance to Innovation ($r^2_{ni} = 4.4\%$) than to the target theme (Mission).
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.13</td>
<td>.09</td>
<td>.01</td>
<td>.15</td>
<td>.17</td>
<td>.03</td>
<td>.08</td>
<td>.01</td>
<td>.05</td>
<td>-.04</td>
<td>.22</td>
<td>.07</td>
</tr>
<tr>
<td>14</td>
<td>.12</td>
<td>.16</td>
<td>.08</td>
<td>.01</td>
<td>.12</td>
<td>.13</td>
<td>-.03</td>
<td>.15</td>
<td>-.01</td>
<td>.02</td>
<td>.12</td>
<td>.02</td>
</tr>
<tr>
<td>26</td>
<td>.01</td>
<td>.06</td>
<td>.12</td>
<td>-.03</td>
<td>-.02</td>
<td>.06</td>
<td>.04</td>
<td>.03</td>
<td>.07</td>
<td>-.08</td>
<td>.07</td>
<td>.18</td>
</tr>
<tr>
<td>38</td>
<td>.16</td>
<td>.14</td>
<td>.18</td>
<td>.16</td>
<td>.05</td>
<td>-.12</td>
<td>.01</td>
<td>.15</td>
<td>.01</td>
<td>-.08</td>
<td>.01</td>
<td>.05</td>
</tr>
<tr>
<td>50</td>
<td>.03</td>
<td>.02</td>
<td>.11</td>
<td>-.04</td>
<td>-.01</td>
<td>-.02</td>
<td>-.07</td>
<td>-.11</td>
<td>.11</td>
<td>.03</td>
<td>.01</td>
<td></td>
</tr>
</tbody>
</table>

Table 19: Analysis of items comprising Empathy.

*Numbers in italics represent individual items comprising a T.P.I. Theme.


Coefficients in bold represent T.P.I. theme part-whole correlations.

Illustrated in Table 17 are the items purporting to measure Empathy (2, 14, 26, 38, and 50), and only items 14 and 38 share a slight amount of variance ($r^2=2.6\%, 2.0\%$ respectively) with this target theme. Also revealed in Table 17 is that item 14 shares less variance with all undesignated themes than with the designated theme (Empathy). All items purporting to comprise Empathy contribute less variance with Investment than with Empathy. Item 50, however, contributes more variance to Mission (0.09\%), Rapport (1.2\%), Individualized Perception (0.16\%), Input Drive (0.50\%), Objectivity (1.2\%), Innovation (1.2\%), and Gestalt (0.09\%) than item 50 contributes to Empathy (0.04\%).
Table 20: Analysis of items comprising Rapport Drive.

*a Numbers in italics represent individual items comprising a T.P.I. Theme.
*c Coefficients in bold represent T.P.I. theme part-whole correlations

As shown in Table 18, when the composite score for Rapport is corrected for an individual item's contribution, through a part-whole correlational technique, only items 3 and 27 contribute meaningful variance (4.8%, 2.9% respectively) to the Rapport Drive theme. In fact, item 3 contributes more variance to Rapport Drive than to any other theme. Item 51, however, reveals a much different pattern in that this item contributes more variance to nine undesignated themes than it does to the designated Rapport Drive theme. Merely one example is how item 51 contributes more variance to Focus (4.4%) than item 51 contributes to Rapport Drive (0.04%).
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-.27</td>
<td>-.01</td>
<td>.11</td>
<td>.05&lt;sup&gt;c&lt;/sup&gt;</td>
<td>.14</td>
<td>.01</td>
<td>-.01</td>
<td>.04</td>
<td>.03</td>
<td>.09</td>
<td>-.10</td>
<td>-.14</td>
</tr>
<tr>
<td>16</td>
<td>.00</td>
<td>.07</td>
<td>.10</td>
<td>.09</td>
<td>.21</td>
<td>.15</td>
<td>.00</td>
<td>.25</td>
<td>.12</td>
<td>.16</td>
<td>-.01</td>
<td>.06</td>
</tr>
<tr>
<td>28</td>
<td>.11</td>
<td>.08</td>
<td>.20</td>
<td>.19</td>
<td>.14</td>
<td>-.08</td>
<td>.12</td>
<td>.27</td>
<td>.12</td>
<td>.16</td>
<td>-.05</td>
<td>.12</td>
</tr>
<tr>
<td>40</td>
<td>.09</td>
<td>.15</td>
<td>.04</td>
<td>.00</td>
<td>.04</td>
<td>.13</td>
<td>.06</td>
<td>-.07</td>
<td>.20</td>
<td>.07</td>
<td>.17</td>
<td>.21</td>
</tr>
<tr>
<td>52</td>
<td>.02</td>
<td>.12</td>
<td>-.01</td>
<td>.10</td>
<td>.10</td>
<td>.05</td>
<td>-.05</td>
<td>-.06</td>
<td>.06</td>
<td>-.02</td>
<td>.08</td>
<td>-.03</td>
</tr>
</tbody>
</table>

Table 21: Analysis of items comprising Individualized Perception.

<sup>a</sup>Numbers in italics represent individual items comprising a T.P.I. Theme.

<sup>b</sup>Roman Numerals indicate T.P.I. Themes: I. Mission, II. Empathy, III. Rapport, IV. Individualized Perception, V. Listening, VI. Investment, VII. Input Drive, VIII. Activation, IX. Objectivity, X. Innovation, XI. Gestalt, XII. Focus.

<sup>c</sup>Coefficients in bold represent T.P.I. theme part-whole correlations

Of the items purporting to measure Individualized Perception (4, 16, 28, 40, 52), only item 28 contributes meaningful variance (3.6%) to the target theme. Without exception, all items were found to share more variance with at least one undesigned theme than the designated theme (Individualized Perception). To illustrate, item 4 shares more variance with Mission (7.3%) than with Individualized Perception (0.26%); items 16 and 28 share more variance with Activation (6.3%, 7.3% respectively) than these items share with the target theme; and item 52 shares more variance with Empathy (1.4%) than with Individualized Perception (1.0%). Also, item 40 contributes more variance to eleven different themes than item 40 contributes to Individualized Perception (.00%).
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-15</td>
<td>-03</td>
<td>.08</td>
<td>.07</td>
<td>.09</td>
<td>.01</td>
<td>.03</td>
<td>.04</td>
<td>.10</td>
<td>.05</td>
<td>.07</td>
<td>.25</td>
</tr>
<tr>
<td>17</td>
<td>-.03</td>
<td>.06</td>
<td>-.01</td>
<td>.18</td>
<td>.11</td>
<td>.06</td>
<td>-.13</td>
<td>.10</td>
<td>.16</td>
<td>.03</td>
<td>-.18</td>
<td>-.11</td>
</tr>
<tr>
<td>29</td>
<td>.20</td>
<td>.24</td>
<td>.25</td>
<td>.11</td>
<td>.15</td>
<td>-.01</td>
<td>.04</td>
<td>.22</td>
<td>.01</td>
<td>.20</td>
<td>.07</td>
<td>-.01</td>
</tr>
<tr>
<td>41</td>
<td>.20</td>
<td>-.03</td>
<td>.22</td>
<td>.11</td>
<td>.11</td>
<td>-.04</td>
<td>.12</td>
<td>.09</td>
<td>.18</td>
<td>-.03</td>
<td>.04</td>
<td>-.05</td>
</tr>
<tr>
<td>53</td>
<td>-.06</td>
<td>.06</td>
<td>.10</td>
<td>.15</td>
<td>.00</td>
<td>-.16</td>
<td>.00</td>
<td>.14</td>
<td>-.01</td>
<td>.01</td>
<td>.06</td>
<td>.01</td>
</tr>
</tbody>
</table>

Table 22: Analysis of items comprising Listening.

- Numbers in italics represent individual items comprising a T.P.I. Theme.
- Coefficients in bold represent T.P.I. theme part-whole correlations.

An examination of Table 20, reveals that when the composite score for Listening is corrected for an individual item’s contribution, through a part-whole correlational technique, only item 29 shares a slight amount of variance (2.3%) with the target theme (Listening). In fact, item 29 shares substantially more variance with Rapport Drive (6.3%) than item 29 shares with Listening (2.3%). Likewise, the variance shared between item 29 and Listening (2.3%), corrected through part-whole correlation, is less than the variance shared between item 29 and Mission (4.0%), Empathy (5.8%), Activation (4.8%), and Innovation (4.0%).

114
<table>
<thead>
<tr>
<th></th>
<th>I⁹</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>-.02</td>
<td>-.09</td>
<td>-.02</td>
<td>-.07</td>
<td>.01</td>
<td>.13</td>
<td>-.06</td>
<td>.01</td>
<td>.08</td>
<td>.14</td>
<td>.13</td>
<td>.21</td>
</tr>
<tr>
<td>18</td>
<td>-.03</td>
<td>.03</td>
<td>.04</td>
<td>-.04</td>
<td>.29</td>
<td>-.13</td>
<td>.06</td>
<td>-.02</td>
<td>.02</td>
<td>.03</td>
<td>.19</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>.11</td>
<td>.03</td>
<td>.03</td>
<td>.12</td>
<td>-.01</td>
<td>.17</td>
<td>.01</td>
<td>.22</td>
<td>.20</td>
<td>.08</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>42</td>
<td>-.20</td>
<td>-.14</td>
<td>-.25</td>
<td>.08</td>
<td>-.18</td>
<td>.27</td>
<td>-.14</td>
<td>.08</td>
<td>-.10</td>
<td>-.10</td>
<td>.00</td>
<td>.07</td>
</tr>
<tr>
<td>54</td>
<td>.02</td>
<td>.20</td>
<td>.15</td>
<td>.10</td>
<td>.08</td>
<td>.20</td>
<td>-.04</td>
<td>.04</td>
<td>.26</td>
<td>.04</td>
<td>-.04</td>
<td>.25</td>
</tr>
</tbody>
</table>

Table 23: Analysis of items comprising Investment.

⁹Numbers in italics represent individual items comprising a T.P.I. Theme.  
¹⁰Coefficients in bold represent T.P.I. theme part-whole correlations

For items purporting to measure Investment (6, 18, 30, 42, and 54), items 18, 42, and 54 share a moderate amount of variance (8.4%, 7.3%, 4.0% respectively) with the target theme (Investment). In fact, items 18 and 42 share less variance with all undesignated themes than these two items share with the designated theme Investment. Item 54, though, shares more variance with Objectivity (6.8%) and Focus (6.3%) than item 54 shares with Investment (4.0%). Also, item 6 contributes more variance to Innovation (2.0%) and Focus (4.4%) than to Investment (1.7%).
Table 24: Analysis of items comprising Input Drive.

*a Numbers in italics represent individual items comprising a T.P.I. Theme.
*c Coefficients in bold represent T.P.I. theme part-whole correlations

Revealed in Table 22 are the items purporting to measure Input Drive (7, 19, 31, 43, and 55), and only item 19 contributes a slight amount of variance (2.6%) to the target theme. All items purporting to comprise Input Drive were found to share more variance with at least one undesignated theme than with the designated theme. For example, item 7 contributes more variance to Mission (3.2%) than it contributes to Input Drive (0.25%), item 19 contributes more variance to Individualized Perception (4.0%) than to Input Drive (2.6%), item 43 shares more variance with Mission (4.8%) than with Input Drive (0.36%), and item 55 shares more variance with Investment (3.2%) than with Input Drive (0.64%).
Table 25: Analysis of items comprising Activation.

\(^a\) Numbers in italics represent individual items comprising a T.P.I. Theme.
\(^b\) Roman Numerals indicate T.P.I. Themes: I. Mission, II. Empathy, III. Rapport, IV. Individualized Perception, V. Listening, VI. Investment, VII. Input Drive, VIII. Activation, IX. Objectivity, X. Innovation, XI. Gestalt, XII. Focus.
\(^c\) Coefficients in bold represent T.P.I. theme part-whole correlations

An examination of Table 23 reveal that when the composite score representing Activation is corrected for an individual item’s contribution, through a part-whole correlational technique, all five items fail to contribute any meaningful variance to the designated theme. The most amount of variance is contributed by item 44, which shares a meager .25 % of the variance associated with Activation. In fact, all items purporting to measure Activation were found to share less variance with the target theme and more variance with all themes other than Activation. Another observation is that all items purporting to comprise Activation contribute more variance to Objectivity and to Focus than to Activation.
<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>11&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.02</td>
<td>.17</td>
<td>.17</td>
<td>.19</td>
<td>.12</td>
<td>.14</td>
<td>-.12</td>
<td>.19</td>
<td>.20&lt;sup&gt;c&lt;/sup&gt;</td>
<td>.05</td>
<td>.07</td>
<td>.09</td>
</tr>
<tr>
<td>23</td>
<td>.08</td>
<td>-.08</td>
<td>-.08</td>
<td>.02</td>
<td>.08</td>
<td>.05</td>
<td>.22</td>
<td>.13</td>
<td>.14</td>
<td>.04</td>
<td>.04</td>
<td>.21</td>
</tr>
<tr>
<td>35</td>
<td>.05</td>
<td>.04</td>
<td>.04</td>
<td>.05</td>
<td>.05</td>
<td>.18</td>
<td>.05</td>
<td>.11</td>
<td>.13</td>
<td>.06</td>
<td>-.05</td>
<td>.06</td>
</tr>
<tr>
<td>47</td>
<td>-.04</td>
<td>-.01</td>
<td>-.01</td>
<td>.03</td>
<td>.15</td>
<td>.04</td>
<td>.05</td>
<td>-.07</td>
<td>-.05</td>
<td>-.04</td>
<td>.02</td>
<td>.06</td>
</tr>
<tr>
<td>59</td>
<td>-.08</td>
<td>-.13</td>
<td>-.13</td>
<td>.23</td>
<td>.05</td>
<td>-.01</td>
<td>.09</td>
<td>.07</td>
<td>.16</td>
<td>-.07</td>
<td>.03</td>
<td>-.09</td>
</tr>
</tbody>
</table>

Table 26: Analysis of items comprising Objectivity.

<sup>a</sup>Numbers in italics represent individual items comprising a T.P.I. Theme.<br/>
<sup>b</sup>Roman Numerals indicate T.P.I. Themes: I. Mission, II. Empathy, III. Rapport, IV. Individualized Perception, V. Listening, VI. Investment, VII. Input Drive, VIII. Activation, IX. Objectivity, X. Innovation, XI. Gestalt, XII. Focus.<br/>
<sup>c</sup>Coefficients in bold represent T.P.I. theme part-whole correlations.

The part-whole correlations illustrated in Table 24 show that of the items purportedly measuring Objectivity (11, 23, 35, 47, and 59), item 11 shares the most amount of variance (4.0%) with the target theme. Furthermore, item 11 shares more variance with the Objectivity theme (4.0%) than with any other theme. Item 23, however, contributes more variance to Input Drive (4.8%) and to Focus (4.4%) than item 23 contributes to Objectivity (2.0%). Also, item 35 shares less variance with Objectivity (1.7%) than item 35 shares with Investment (2.3%).
<table>
<thead>
<tr>
<th>Item</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>-.09</td>
<td>-.10</td>
<td>-.01</td>
<td>-.08</td>
<td>.04</td>
<td>.15</td>
<td>-.04</td>
<td>-.02</td>
<td>-.03</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>-.05</td>
<td>-.02</td>
<td>.20</td>
<td>.08</td>
<td>.14</td>
<td>-.15</td>
<td>.11</td>
<td>.17</td>
<td>.04</td>
<td>-.03</td>
<td>-.05</td>
<td>-.12</td>
</tr>
<tr>
<td>33</td>
<td>-.05</td>
<td>-.10</td>
<td>.10</td>
<td>.11</td>
<td>-.07</td>
<td>.06</td>
<td>.04</td>
<td>.03</td>
<td>.13</td>
<td>-.01</td>
<td>-.06</td>
<td>-.08</td>
</tr>
<tr>
<td>45</td>
<td>-.02</td>
<td>-.14</td>
<td>.04</td>
<td>.05</td>
<td>.01</td>
<td>.12</td>
<td>-.08</td>
<td>.15</td>
<td>.01</td>
<td>.03</td>
<td>.04</td>
<td>-.07</td>
</tr>
<tr>
<td>57</td>
<td>-.08</td>
<td>.18</td>
<td>.08</td>
<td>.12</td>
<td>.07</td>
<td>.00</td>
<td>-.02</td>
<td>.04</td>
<td>-.08</td>
<td>-.10</td>
<td>.14</td>
<td>.05</td>
</tr>
</tbody>
</table>

Table 27: Analysis of items comprising Innovation.

*Numbers in italics represent individual items comprising a T.P.I. Theme.
*cCoefficients in bold represent T.P.I. theme part-whole correlations

According to Table 25, when the Innovation composite score is corrected for an individual item’s contribution to the composite, two items (9, 57) contribute the same amount of variance (1.0%) to the Innovation theme. In general, the five items purporting to measure Innovation (9, 21, 33, 45, and 57) fail to contribute meaningful variance to the target theme. In fact, item 33 contributes more variance to the eleven themes other than item 33 contributes to Innovation, while item 21 contributes more variance to ten themes other than Innovation.
<table>
<thead>
<tr>
<th></th>
<th>I⁵</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>.08</td>
<td>.06</td>
<td>.07</td>
<td>.00</td>
<td>.01</td>
<td>.04</td>
<td>.16</td>
<td>.10</td>
<td>-.02</td>
<td>-.13</td>
<td>-.13</td>
<td>-.01</td>
</tr>
<tr>
<td>22</td>
<td>.01</td>
<td>.02</td>
<td>.11</td>
<td>-.10</td>
<td>.05</td>
<td>-.08</td>
<td>-.13</td>
<td>.05</td>
<td>.06</td>
<td>.14</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>-.13</td>
<td>.01</td>
<td>-.11</td>
<td>-.03</td>
<td>-.08</td>
<td>.02</td>
<td>-.19</td>
<td>-.06</td>
<td>-.09</td>
<td>.06</td>
<td>.14</td>
<td>.05</td>
</tr>
<tr>
<td>46</td>
<td>.11</td>
<td>.20</td>
<td>.02</td>
<td>.06</td>
<td>.00</td>
<td>.09</td>
<td>-.12</td>
<td>-.02</td>
<td>.10</td>
<td>.09</td>
<td>.13</td>
<td>.04</td>
</tr>
<tr>
<td>58</td>
<td>.07</td>
<td>.03</td>
<td>-.10</td>
<td>.14</td>
<td>.10</td>
<td>-.07</td>
<td>.01</td>
<td>-.02</td>
<td>.03</td>
<td>.00</td>
<td>-.08</td>
<td></td>
</tr>
</tbody>
</table>

Table 28: Analysis of items comprising Gestalt.

⁵Numbers in italics represent individual items comprising a T.P.I. Theme.
⁷Coefficients in bold represent T.P.I. theme part-whole correlations.

Part-whole correlations computed for the Gestalt composite reveal that items 10 and 46 both contribute the same amount of variance (1.7%) to the target theme while items 22 and 34 contribute also the same amount of variance (2.0%) to Gestalt. Item 58, however, fails to contribute any variance (0.0%) to the Gestalt theme. Furthermore, all themes share more variance with item 58 than this item shares with Gestalt. Items 10, 22, and 34, however, share more variance with Gestalt than with all other themes.
Table 29: Analysis of items comprising Focus.

\[ \begin{array}{cccccccccccc}
 & I^b & II & III & IV & V & VI & VII & VIII & IX & X & XI & XII \\
12^a & .12 & .11 & .07 & -.09 & -.05 & .13 & -.01 & .02 & -.04 & .03 & .12 & .16^e \\
24 & .02 & .07 & .11 & .28 & .11 & .20 & .07 & .03 & .37 & .15 & -.02 & .03 \\
36 & .16 & .10 & -.01 & -.01 & -.06 & .11 & .09 & .01 & .01 & -.11 & .10 & .13 \\
48 & -.03 & .11 & .19 & -.02 & -.03 & .03 & .06 & .05 & -.04 & -.06 & .03 & .04 \\
60 & .03 & -.07 & .07 & .03 & .13 & .21 & .05 & -.02 & .00 & -.08 & -.10 & .23 \\
\end{array} \]

\*Numbers in italics represent individual items comprising a T.P.I. Theme.
\*Coefficients in bold represent T.P.I. theme part-whole correlations

For items purporting to measure Focus (12, 24, 36, 48, and 60), item 60 shares the most variance (5.3%) with the Focus theme. Items 12 and 60 contribute more variance to Focus (2.6% and 5.3% respectively) than these items contribute to any other theme. Item 24, though, shares more variance with Empathy (.50%), Rapport (1.2%), Individualized Perception (7.8%), Listening (1.2%), Investment (4.0%), Input Drive (.50%), Objectivity (13.7%), and Innovation (2.3%) than item 24 shares with Focus. One theme, Gestalt, shares less variance with all items purporting to measure Focus than all items share with Focus.
As in Study I, the two criterion variables chosen for Study II are (1) ratings provided by each teacher's evaluating principal and (2) the number of days the teacher missed work. To assess the relationship between interview scores and the rating variable, a correlational analysis was conducted. As illustrated in Table 30, the highest degree of relationship between predicted score and observed score is the 3.2% of the variance shared between the principals' rating of both Individualized Perception and of Gestalt and the predictor scores attained for Individualized Perception and Gestalt. Revealed also in Table 30 is that the highest degree of relationship between a predictor variable and the absenteeism variable is the 3.2 percent of variance shared between absenteeism and predicted scores attained for both Listening and for Gestalt.
<table>
<thead>
<tr>
<th></th>
<th>R* I</th>
<th>R* II</th>
<th>R* III</th>
<th>R* IV</th>
<th>R* V</th>
<th>R* VI</th>
<th>R* VII</th>
<th>R* VIII</th>
<th>R* IX</th>
<th>R* X</th>
<th>R* XI</th>
<th>R* XII</th>
<th>Absenteeism</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Mission (5)b</td>
<td>.17</td>
<td>-.05</td>
<td>.04</td>
<td>-.05</td>
<td>-.02</td>
<td>-.03</td>
<td>-.06</td>
<td>-.02</td>
<td>-.04</td>
<td>-.03</td>
<td>-.11</td>
<td>-.11</td>
<td>.02</td>
</tr>
<tr>
<td>II. Empathy (5)</td>
<td>.16</td>
<td>.15</td>
<td>.02</td>
<td>.13</td>
<td>.17</td>
<td>.18</td>
<td>.05</td>
<td>.03</td>
<td>.07</td>
<td>.09</td>
<td>.05</td>
<td>.14</td>
<td>-.09</td>
</tr>
<tr>
<td>III. Rapport (5)</td>
<td>.17</td>
<td>.19</td>
<td>.04</td>
<td>.16</td>
<td>.17</td>
<td>.26</td>
<td>.30</td>
<td>.12</td>
<td>.11</td>
<td>.15</td>
<td>.11</td>
<td>.23</td>
<td>-.17</td>
</tr>
<tr>
<td>IV. Individualized Perception (5)</td>
<td>.00</td>
<td>.17</td>
<td>.16</td>
<td>.18</td>
<td>.05</td>
<td>.07</td>
<td>.22</td>
<td>.05</td>
<td>.15</td>
<td>.07</td>
<td>.16</td>
<td>.13</td>
<td>-.10</td>
</tr>
<tr>
<td>V. Listening (5)</td>
<td>.07</td>
<td>-.04</td>
<td>-.10</td>
<td>.01</td>
<td>.03</td>
<td>.03</td>
<td>-.02</td>
<td>-.09</td>
<td>-.06</td>
<td>-.04</td>
<td>-.08</td>
<td>.01</td>
<td>-.18</td>
</tr>
<tr>
<td>VI. Investment (5)</td>
<td>.09</td>
<td>.09</td>
<td>.05</td>
<td>.11</td>
<td>.12</td>
<td>.11</td>
<td>.13</td>
<td>.08</td>
<td>.05</td>
<td>.14</td>
<td>.22</td>
<td>.09</td>
<td>-.17</td>
</tr>
<tr>
<td>VII. Input Drive (5)</td>
<td>-.01</td>
<td>-.01</td>
<td>-.11</td>
<td>-.01</td>
<td>.03</td>
<td>.02</td>
<td>.02</td>
<td>-.03</td>
<td>-.12</td>
<td>.06</td>
<td>-.01</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>VIII. Activation (5)</td>
<td>.12</td>
<td>.01</td>
<td>.00</td>
<td>-.07</td>
<td>.03</td>
<td>.04</td>
<td>.11</td>
<td>.04</td>
<td>.06</td>
<td>.00</td>
<td>.14</td>
<td>.11</td>
<td>-.08</td>
</tr>
<tr>
<td>IX. Objectivity (5)</td>
<td>.15</td>
<td>.12</td>
<td>.07</td>
<td>.15</td>
<td>.07</td>
<td>.11</td>
<td>.13</td>
<td>.12</td>
<td>.14</td>
<td>.13</td>
<td>.14</td>
<td>.10</td>
<td>-.17</td>
</tr>
<tr>
<td>X. Innovation (5)</td>
<td>-.06</td>
<td>.07</td>
<td>.04</td>
<td>-.08</td>
<td>-.05</td>
<td>-.06</td>
<td>.05</td>
<td>-.01</td>
<td>.05</td>
<td>.08</td>
<td>.08</td>
<td>.05</td>
<td>.07</td>
</tr>
<tr>
<td>XI. Gestalt (5)</td>
<td>.09</td>
<td>.25</td>
<td>.19</td>
<td>.20</td>
<td>.23</td>
<td>.23</td>
<td>.16</td>
<td>.06</td>
<td>.14</td>
<td>.17</td>
<td>.18</td>
<td>.20</td>
<td>-.18</td>
</tr>
<tr>
<td>XII. Focus (5)</td>
<td>-.04</td>
<td>.02</td>
<td>-.10</td>
<td>-.03</td>
<td>.07</td>
<td>-.01</td>
<td>.25</td>
<td>.00</td>
<td>-.03</td>
<td>.07</td>
<td>.22</td>
<td>-.03</td>
<td>-.10</td>
</tr>
<tr>
<td>Absenteeism</td>
<td>-.08</td>
<td>-.05</td>
<td>-.10</td>
<td>-.05</td>
<td>-.02</td>
<td>-.14</td>
<td>-.14</td>
<td>-.05</td>
<td>-.11</td>
<td>-.10</td>
<td>-.20</td>
<td>-.02</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 30: Relationship between predictor (T.P.I.) and criterion measures (administrative ratings and absenteeism)

*a Bolded column headings represent administrator ratings of teaching performance by T.P.I. theme.

b Number in parentheses indicates the number of items contained within each T.P.I. theme.
To determine if a linear combination of predictor variables could account for a meaningful amount of variance associated with absenteeism, a regression analysis was conducted. This analysis utilizes a full model with all twelve 12 themes regressed against the criterion variable absenteeism using a blockwise entry. Contained in Table 32 is the omnibus test of the regression model.

Collectively, these 12 themes account for 14% of the variance associated with absenteeism. The individual contributions of each theme, given other themes in the model, are found in Table 32. The most important themes of this particular linear combination are Investment and Gestalt, while the least important theme for this linear combination is Activation.

An examination of the direction of loadings (see Standardized Beta Coefficients in Table 32) reveals some interesting findings. High scores on Innovation are likely to be associated with high absenteeism rates. This may suggest that an individual concerned with discovering new and creative ideas, values novelty more than a routine work environment. That is, individuals seeking creativity may choose to fulfill this type of need by missing work, which may actually constrict creative thinking.

This interpretation is somewhat supported by the negative relationship between Gestalt and absenteeism (see Standardized Beta Coefficients in Table 32). Teachers that have a drive toward completeness and teachers who feel uneasy until work is finished (high Gestalt) are more likely to attend work than those teachers with low Gestalt. Note also that Activation fails to account for any variance, individually or collectively, shared with absenteeism.
<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>89.49</td>
<td>12</td>
<td>7.46</td>
<td>1.50</td>
<td>.135</td>
</tr>
<tr>
<td>Residual</td>
<td>552.22</td>
<td>111</td>
<td>5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>641.71</td>
<td>123</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 31: Full Regression model containing all themes.

<table>
<thead>
<tr>
<th>Variables</th>
<th>b</th>
<th>Beta</th>
<th>T Test</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission</td>
<td>.12</td>
<td>.06</td>
<td>3.94</td>
<td>.56</td>
</tr>
<tr>
<td>Investment</td>
<td>-.32</td>
<td>-.16</td>
<td>.59</td>
<td>.11</td>
</tr>
<tr>
<td>Focus</td>
<td>-4.19E-02</td>
<td>-.02</td>
<td>-1.62</td>
<td>.84</td>
</tr>
<tr>
<td>Objectivity</td>
<td>-.13</td>
<td>-.07</td>
<td>-.21</td>
<td>.50</td>
</tr>
<tr>
<td>Listening</td>
<td>-.28</td>
<td>-.15</td>
<td>-.68</td>
<td>.13</td>
</tr>
<tr>
<td>Empathy</td>
<td>-2.96E-02</td>
<td>-.02</td>
<td>-1.54</td>
<td>.87</td>
</tr>
<tr>
<td>Rapport</td>
<td>-.27</td>
<td>-.15</td>
<td>-.16</td>
<td>.15</td>
</tr>
<tr>
<td>Activation</td>
<td>9.68E-03</td>
<td>.05</td>
<td>-1.44</td>
<td>.96</td>
</tr>
<tr>
<td>Innovation</td>
<td>.32</td>
<td>.14</td>
<td>.05</td>
<td>.14</td>
</tr>
<tr>
<td>Gestalt</td>
<td>-.25</td>
<td>-.16</td>
<td>1.48</td>
<td>.07</td>
</tr>
<tr>
<td>Input Dive</td>
<td>.15</td>
<td>.07</td>
<td>-1.78</td>
<td>.49</td>
</tr>
<tr>
<td>Individualized Perception</td>
<td>-3.65</td>
<td>-.02</td>
<td>.69</td>
<td>.84</td>
</tr>
<tr>
<td>Constant</td>
<td>5.81</td>
<td>-.11</td>
<td></td>
<td>.00</td>
</tr>
</tbody>
</table>

Table 32: Theme contribution to regression model.

Note: $R^2 = .37$

$R^2 = .14$
Page 126 is missing
CHAPTER 5
FINDINGS, IMPLICATIONS, LIMITATIONS

The employment interview has maintained an important and frequent role as a predictor of job performance. Several researchers indicate that most all people seeking employment are likely to experience some form of face to face interview prior to being hired (Campion et al., 1997; Dipboye, 1992; Weston & Warmke, 1988). Not surprising, then, is that research on employee selection has been conducted for nearly 80 years (Campion et al., 1997).

Much of the early selection research, conducted in the private sector, involved an approach which focused on assessing the reliability and the validity of the employment interview as a predictor of subsequent job performance (Mayfield, 1964; Ulrich & Trumbo, 1965; Wagner, 1949; Wright, 1969). One of the fundamental components of this approach, as suggested by Wagner (1949), is that structured interview formats be utilized. This early research approach concerning the interview as a predictor of job performance was described by Mayfield as “macroanalytic.”

Because, in many instances, the macroanalytic approach failed to establish reliability and validity for the interview, a different research approach was considered. This alternative selection research approach involved a “microanalytic” approach and
focused on identifying specific factors influencing decision making within the selection process (Mayfield, 1964). Although the microanalytic approach to selection has improved the reliability and the validity of the selection process, macroanalytic techniques remain viable in situations where a pre-employment test is also an interview.

Even with selection research in the private sector being conducted for nearly 80 years, public sector research on teacher selection has not experienced the same depth of inquiry. Administrators and researchers interested in examining the teacher selection process have relied on research from the private sector to provide an historical perspective of the research literature. Although Wagner (1949) suggested the standardization of personnel selection, few school districts are currently attempting to validate or standardize the selection process.

Employment decisions in education, unlike the private sector, are difficult in part because performance standards are often difficult to measure. These difficulties tend to exist because there often fails to be consensus regarding a definition of effective teaching or an understanding of how to measure performance standards reliably. In one the most comprehensive examinations of teacher selection, Schalock (1979) stated, “A teacher’s knowledge of either content or method will never be sufficient as a predictor of teaching success . . .” (p.376).

It should not be surprising, then, given the importance of making wise teacher selection decisions, that a foundation of empirical teacher selection research has emerged (see Bolton, 1969; Schalock, 1979; Young & Ryerson, 1986; Young & Heneman, 1986;
Young et. al., 1989; Young et. al., 1993). This stream of research serves as the cornerstone for current and for future teacher selection research efforts.

Utilizing the established foundation of selection research, from both the private sector and from the educational arena, the present study attempts to examine the extent to which predictive relationships may exist between the employment interview and job performance. If, according to Schalock, neither content nor method can predict teaching “success,” then perhaps behavior or personality characteristics can be used as predictors of teacher “success.” One such predictive measure, increasingly utilized in school systems, is the commercially designed Teacher Perceiver Interview (T.P.I.), both abbreviated and complete forms.

The Teacher Perceiver Interview purports to predict future teaching performance according to the following twelve themes: (1) Mission, (2) Empathy, (3) Rapport Drive, (4) Individualized Perception, (5) Listening, (6) Investment, (7) Input Drive, (8) Activation, (9) Innovation, (10) Gestalt, (11) Objectivity, and (12) Focus. These themes are derived from the research conducted by Dr. Donald Clifton who, shortly after Wagner’s (1949) review of selection research, examined the behavior and personality characteristics of successful college counselors. The development of the Teacher Perceiver Interview instrument, in 1979, is an attempt by the testing company to identify “. . . the most accurate information about a person’s potential job performance in the least amount of time” (Gallup Organization, 1993, p.2).

To address the research questions posited in Chapter One, two studies were conducted. Study I consists of 72 classroom teachers in a large school district in the
southeastern United States who were administered the abbreviated form (twenty-two questions) of the T.P.I. Study II consists of 124 classroom teachers in a large school system in the mid-western United States who were administered the complete, sixty-question, T.P.I. In both studies, the predictor variable is the scores achieved on the T.P.I. instrument and the criterion variables are (1) teacher performance ratings provided by building level principals and (2) the number of days the participating teachers missed work.

Predictor scores are derived by summing subject responses to either the abbreviated or the complete version of the T.P.I. Responses to open-ended questions were evaluated by the Director of Personnel, in Study I, and by the Coordinator of Human Resources, in Study II, who coded “listen for” responses from subjects. “Listen for” responses are those responses, provided by interviewees, that contain specific words or phrases for which the “perceiver” is trained to recognize. Training in the coding of “listen for” responses is both required by and provided from the testing company.

To achieve status as a “perceiver,” both the Director of Personnel and the Coordinator of Human Resources received certification that required over 100 hours of study and instruction in the administration of the Teacher Perceiver Interview. This training focused on assisting “perceivers” in coding accurately “listen for” responses. Furthermore, the “perceiver” must achieve 85% consistency on item by item T.P.I. scoring when compared to Selection Research Incorporated (Gallup, Inc.) staff members (Schmidt & Rader, 1999).
Mission with respect to the complete version, fared substantially better than the abbreviated version relative to internal consistency (.04 versus .55). However, within the complete version, only items 1, 37, and 49 correlated substantially more with Mission than with other themes. Although the Mission theme, as defined by the complete version, correlated the same as the Mission theme as defined by the abbreviated version, for absenteeism (.02 in both instances), the correlation between the complete and the abbreviated version with principals' ratings was only slightly better for the former (.12 and .17, respectively).

Collectively, the theme of Mission, as assessed by either the abbreviated or the complete version, fails to account for substantial variance associated with absenteeism individually or in combination with other variables (see Betas) in a linear combination. Mission, however, was independent of all other themes except Individualized Perception in the abbreviated model (-.20) and Input Drive in the complete model (.22).

**Empathy**

Reliability for Empathy was non-existent for both the abbreviated and complete versions (.00 and .01, respectively). With respect to the abbreviated version, the items chosen to measure Empathy (26, 38, and 50) contribute very little variance to the target theme. Furthermore, these items correlate higher with other themes (Item 26 with 5 other themes, Item 38 with 3 other themes, and Item 50 with 1 other theme) than with the Empathy theme. Selecting alternative items from the complete version theme for Empathy, however, fails to contribute additional meaningful variance. Also, Empathy as
defined by the abbreviated version, contributed less than 3% of the variance to
absenteeism.

With respect to the complete version of the Empathy theme (items 2, 14, 26, 38
and 50), internal consistency was hardly noticeable (.01). In fact, four of these items
correlated higher with other themes (Item 2 with 4 themes, Item 26 with 5 themes, Item
38 with 3 themes, and Item 50 with 6 themes) than with the Empathy theme. Only item
14 correlated higher with Empathy than any other theme. Further examination also
revealed that Empathy correlates somewhat more with principals’ ratings in the complete
version (.15) than with the abbreviated version (.04).

Empathy, as assessed by the complete version, contributes little meaningful
variance to absenteeism, individually or in combination with other variables. The
abbreviated version, however, contributes less variance to absenteeism independently
(2.9) than it contributes when placed in combination with other variables (see Betas).

**Rapport**

In reference to Rapport, reliability was non-existent for the abbreviated
version (.00) and low for the complete version (.21). The items chosen for the abbreviated
version of Rapport (3,39) demonstrate low part-whole correlations (.09 for both), and
these items correlated higher with themes other than with Rapport (Item 3 with 6 other
themes and Item 39 with 2 other themes). Item 27, however, would have been a better
item for the abbreviated version than item 39 because the corrected part-whole
correlation for item 27 is (.17) as opposed to -.09 for item 39. Rapport also correlated moderately low with absenteeism (.23).

Rapport, with respect to the complete version, faired somewhat better than the abbreviated version's in relation to internal consistency (.00 versus .21). However, with the exception of Item 3, all items purporting to measure the complete version of the Rapport theme correlate higher with other themes than with the target theme (Item 15 with 4 other themes, Item 27 with 1 other theme, Item 39 with 8 other themes, and Item 51 with 9 other themes). Principal ratings correlated negatively with Rapport (-.18) for the abbreviated version and correlated positively (.04) with the complete version. Collectively, the Rapport theme as assessed by either the abbreviated or the complete versions, account for approximately the same variance with absenteeism, whether individually or in combination with other variables (see Betas).

**Individualized Perception**

The Individualized Perception theme for the abbreviated version demonstrated virtually zero reliability (0.2), and the complete version demonstrated very low reliability (.20). The actual items selected for the abbreviated version (28, 4), appear to be poor choices for several reasons beyond reliability. First, these items correlated higher with other themes (Item 28 with 9 other themes and Item 4 with 8 other themes) than with the target theme. Second, Items 16 and 52, from the complete version, share more variance with the target theme (.81% and 1.00%, respectively).
As for the complete version, all items purporting to measure Individual Perception fail generally to contribute any meaningful variance to the target theme. All items (4, 16, 28, 40 and 52) correlate higher with other themes (Item 4 with 6 other themes, Item 16 with 6 other themes, Item 28 with 2 other themes, Item 40 with 11 other themes, and Item 52 with 1 other theme) than the target theme. Although the Individualized Perception theme correlated the same in the abbreviated version (.01 in both instances) the correlation between the abbreviated and complete versions with principals’ ratings was somewhat larger for the complete version (.18) than for the abbreviated version (-.02). The Individual Perception theme as assessed by either the abbreviated or the complete version fails to contribute any meaningful variance to absenteeism individually or combined (see Betas) in a linear combination.

**Listening**

Reliability for Listening measured only by the complete version, was .27. All items purporting to measure Listening (5, 17, 29, 41, 53) all failed to contribute any meaningful variance to the target theme. In fact, item 53 contributed .00% of the variance to Listening. Furthermore, all items correlated higher with other themes (Item 5 with 3 themes, Item 17 with 4 themes, Item 29 with 5 themes, Item 41 with 4 themes, and Item 53 with 10 themes).

Absenteeism had a negative correlation with Listening (-.18) and when combined with other variables, Listening had a slightly smaller negative correlation (-.15). Ratings
provided by principals' for Listening correlated extremely low with the predicted scores for Listening (.03).

**Input Drive**

Input Drive displayed moderately low reliability for the abbreviated version (.40) and zero reliability (.00) for the complete version. By all accounts though, the items selected to measure Input Drive by the abbreviated version (7, 31, 43, 55) exhibit higher part-whole correlations (.30, .23, .13, and .27 respectively) than the items on the complete version (-.05, -.16, .00, .06, .08). In fact, all items purporting to measure Input Drive for the abbreviated version correlate higher with the target theme than any other theme.

Measuring Input Drive using the complete version appears to be a poor choice. Interestingly, the same 4 items measuring Input Drive (7, 31, 43, and 55) on the abbreviated version failed to contribute any meaningful variance on the full version. The one additional item (19) measure by the full version still managed to correlate higher with 3 themes other than Input Drive. Input Drive accounted for very little variance in Absenteeism using either the abbreviated or complete version (-.08 and .09, respectively).

Input Drive also accounts for little variance associated with Absenteeism individually (-.08 for the screener and .09 for complete) or in combination with other variables (see Betas). Furthermore, the correlation between abbreviated and complete versions is only slightly higher with the abbreviated (.10) than with complete (.02).
Activation

Reliability for Activation was non-existent for the abbreviated version (.00) and only moderately low for the full version (.47). Neither the abbreviated version nor the complete version contained items that contributed any meaningful variance to the target theme (Activation). In fact, items associated with the abbreviated version (20, 32, and 56) correlated higher with other themes (Item 20 with 6 other themes, Item 32 with 3 other themes, and Item 56 with 7 other themes).

The same items using the complete version correlated higher also with other themes (Item 29 with 9 themes, Item 32 with 8 themes, and Item 56 with 9 themes). Items 8 and 44, not measured by the abbreviated version, correlated higher with other themes as well (11 themes and 5 themes respectively).

The correlation between the abbreviated and complete versions with principals’ ratings was higher with the former (.29 and .04, respectively). An examination of absenteeism reveals that only slightly more variance in absenteeism is accounted for using the abbreviated version as opposed to the complete version (−.14 versus −.08, respectively).
Gestalt

Reliability for Gestalt was quite low for the abbreviated version (.14) and equally as low for the complete version (.17). The part-whole correlation for both items comprising the abbreviated version is .07. Furthermore, each item (58 and 22) contributes more variance to themes other than to the target theme (Item 58 with 3 other themes and Item 22 with 2 other themes). With the exception of item 58 from the complete version, any item from the complete version would have been measured better on the abbreviated version because each of the four items contribute more variance to the target theme (.13, .14, .14, .13, respectively).

Although, items comprising the complete version contribute more variance to Gestalt, these part-whole correlations are quite low. Only one item (22) on the complete version fails to correlate higher with themes other than the target theme. The remaining items correlate higher with at least one other theme.

Gestalt correlated somewhat higher for principal ratings on the abbreviated version (.24) than on the complete version (.18). Absenteeism showed a larger negative correlation (.18) for the complete version measuring Gestalt than the abbreviated version measuring Gestalt (.02).

Innovation

The reliability for Innovation was small for the abbreviated version (.27) and slightly higher for the complete version (.35). Neither item purporting to measure Innovation for the abbreviated version (45, 9) contributed any meaningful variance to the
target theme \( (r^2 = 2.6\% \text{ for both}) \). Furthermore, both items contributed more variance to two themes other than the target theme. Both items did, however, account for more variance on the abbreviated version than on the complete version.

Each item comprising the Innovation theme for the complete version (9, 21, 33, 45, and 57) contributed little meaningful variance to the target theme. All items correlated higher with other themes (Item 9 with 2 other themes, Item 21 with 10 other themes, Item 33 with 11 other themes, Item 45 with 8 other themes, and Item 57 with 3 other themes) than the target theme.

Although the Innovation theme, as defined by the compete version correlated about the same for absenteeism as the abbreviated version (.07 and .04, respectively), the correlations between the complete and the abbreviated versions with principals’ ratings was better for the former (.25 and .08 respectively). More variance in absenteeism was accounted for when Innovation for both abbreviated and complete versions was combined with other themes (see Betas).

**Investment**

The reliability for the Investment theme, as measured by the complete version, was a moderate .66. Items comprising Investment (6, 18, 30, 42, 54) each revealed a part-whole correlation of .13, .29, .17, .27, and .20 respectively. Only two items (18, 42) failed to correlate higher with themes other than the target theme. Investment accounted for approximately the same amount of variance associated with absenteeism.
independently and combined with other themes (see Betas). Principals’ ratings of Investment correlated to .11.

Focus

Measured only by the complete version, reliability for Focus was .40. Item 60 contributed the most variance to the target theme ($r^2 = 5.3\%$) while item 24 contributed the least ($r^2 = 0.09\%$). Item 24 also correlated higher with 8 themes other than the target theme (Focus). Items 60 and 12, however, failed to correlate higher with any theme other than the target theme. Focus correlated slightly more negative independently (-.10) than combined with other themes (-.02). There was essentially no relationship between principals’ ratings of Focus and predictor scores for the Focus theme.

Objectivity

Reliability for Objectivity, as measured by the complete version, is .32. Part-whole correlations reflect that item 11 contributes the most variance ($r^2 = 4.0\%$) associated with the target theme and that Item 47 contributes the least ($r^2 = .25\%$). Only Item 11 fails to correlate higher with themes other than Objectivity. All remaining items correlate higher with other themes (Item 23 with 2 other themes, Item 35 with 1, Item 47 with 3 and Item 59 with 1). Objectivity correlated with absenteeism slightly more negatively independent than with combined with other themes (- .17 versus -.07 respectively).
In reference to findings related to the four research questions posed in Chapter I, the following observations are made:

Q1: Little variance, in general, is shared between any of the ten-predictor themes from the T.P.I.S. and the performance ratings provided by building principals. The highest degree of relationship between these two variables is the 8.4% of variance shared between the principals’ ratings of Activation performance and the predictor scores attained for the Activation Theme.

Q2: In general, only a slight amount of variance is shared between any of the ten-predictor themes measured by the T.P.I.S. and absenteeism. The highest degree of relationship between these two variables is the 5.3 percent of variance shared between absenteeism and predictor scores for the Rapport theme.

Q3: Very little variance, in general, is shared between any of the twelve themes measured by the T.P.I. and the performance ratings provided by building principals. The highest degree of relationship between these two variables is the 4.8% of variance shared between the principals’ rating of the Focus theme and the predictor scores attained for Focus.

Q4: In general, very little variance is shared between any of the twelve-predictor themes measured by the T.P.I. and absenteeism. The highest degree of relationship between
these two variables is the 3.2% of variance shared between absenteeism and predictor scores for both the Listening theme and the Gestalt theme.

These findings fail to be surprising given the low reliabilities derived from the theme analysis. Because reliability sets the upper limit for validity, the predictive value of the T.P.I. is of little practical meaning. The lack of internal consistency for the Teacher Perceiver Interview, therefore, has emerged as a significant component to this study.

Implications

There is little doubt that school administrators continue to search for improved interviewing techniques that might better select the most effective classroom teachers. Although the Teacher Perceiver Interview has been used by many school districts nationwide, both the construct validity and the predictive validity of the instrument need further development if this selection system is to provide practical results. One of the implications here is that, if school administrators feel compelled to utilize a commercially designed pre-employment interview instrument, administrators must determine to what extent the instrument measures what it purports to measure.

If, for instance, administrators are concerned about hiring classroom teachers who care greatly about their students, using the Empathy theme as measured by the Teacher Perceiver Interview may lead to an erroneous conclusion. Because items comprising Empathy correlate higher with T.P.I. themes other than with Empathy, there exists a glaring vacancy in the ability of this theme to actually predict what might be considered
empathetic behavior. Moreover, a more reliable predictor of Empathy, as illustrated in Table 7, is the Rapport theme.

Of particular importance to practicing school administrators is that deliberate attempts must be made to select classroom teachers utilizing practices proven reliable and valid. The use of a pre-employment test, to help identify teachers with the necessary qualities to help students learn, can be a useful tool if it is used appropriately and in conjunction with an appropriate evaluation instrument. Just because an instrument is produced commercially and used frequently, however, is not sufficient for automatic adoption by a school district. In short, all districts should conduct validity assessments for selection practices.
Limitations

The present study, like all studies, is effected by some limitations. One limitation of this study applies to the generalizability of results. Specifically, the two school districts participating in this study were chosen because each district administers the Teacher Perceiver Interview. Although the districts chosen are from two different states, the results obtained are generalizable to only the districts represented in this study.

Another limitation to the present study is that no mechanism is in place to control for the possibility that some teacher candidates may have had prior interview experience specific to the Teacher Perceiver instrument. Some variance on the predictor measure may, therefore, be due to a teacher’s prior exposure to the interview questions. Prior exposure to interview questions may have given some interviewees an opportunity to plan a desired response or to anticipate when a specific interview question might be posed.

Yet another limitation to this study is the elapsed time between the predictor assessment and the criterion assessment. The longer a classroom teacher has been teaching, the more time the teacher has to improve teaching performance. For some teachers in this study, the time lapse between pre-employment testing and post-employment ratings was as short as one year and as long as three years.
Conclusion

The continuous improvement of the selection process is a goal that all organizations must strive to achieve. Of importance to those charged with selecting personnel is that the reliability and the validity of all phases of the selection process must be proven and maintained. This study sought specifically to identify and to explain the importance of demonstrating the predictive validity of the selection interview.

The results of this study, although not encouraging for the school districts selected for this study, raise important issues related to the use of any commercially designed selection instrument. First, any interview system, or test, can only be expected to be as reliable and as valid as the items comprising that test. Second, given the costs, both in terms of time and money, school districts must give careful consideration as to the reasons a commercially designed selection instrument is more reliable and more valid than other, more traditional predictors of job performance.

There is clearly a need for further study of the Teacher Perceiver Interview as it relates both to test construction and predictive worth. Recommendations for further research consideration are as follow:

1. It is recommended that the Teacher Perceiver Interview, both the abbreviated version and complete interview, be used in conjunction with other predictive sources of relevant employment information when selecting classroom teachers.
2. The use of this type of commercially designed selection instrument must be tested so as to reveal any possible adverse impact against protected class persons.

3. A comparison study is suggested using the same basic design as the present study using only teachers absent any teaching experience other than student teaching.

4. Factor analysis should be conducted to assess the extent to which there exists twelve distinct dimensions of teaching behavior or if these themes can be otherwise grouped.

5. All school districts utilizing the Teacher Perceiver Interview must conduct validity assessments of the instrument rather than relying on validity studies generalizable only to specific populations.
APPENDIX A

GLOSSARY OF TEACHER PERCEIVER THEMES
Appendix A

Teacher Perceiver Themes

Mission – Mission is what takes some individuals and groups out of society’s mainstream in order to assure the quality and purposiveness of that mainstream. Mission is a deep underlying belief that students can grow and attain self-actualization. A teacher with mission has a goal to make a significant contribution to other people.

Empathy – Empathy is the apprehension and acceptance of the state of mind of another person. Practically, we say we put ourselves into the other person’s place. Empathy is the phenomenon that provides the teacher feedback about the individual student’s feeling and thoughts.

Rapport Drive – The rapport drive is evidenced by the teacher’s ability to have an approving and mutually favorable relationship with each student. The teacher likes students and expects them to reciprocate. Rapport is seen by the teacher as a favorable and necessary condition of learning.
**Individualized Perception** – Individualized perception means that the teacher spontaneously thinks about the interests and needs of each student and makes every effort to personalize each student’s program.

**Listening** – The listening theme is evident when a person spontaneously listens to others with responsiveness and acceptance. Listening is viewed as beneficial to the speaker.

**Investment** – The investment theme is indicated by the teacher’s capacity to receive satisfaction from the growth of students. This is in contrast to the person who must personally perform to achieve satisfaction.

**Input Drive** – Input drive is evidenced by the teacher who is continuously searching for ideas, materials, and experiences to use in helping other people, especially students.

**Activation** – Activation indicates that the teacher is capable of stimulating students to think, to respond, to feel, and to learn.

**Innovation** – The innovation theme is indicated when a teacher tries new ideas and techniques. A certain amount of determination is observed in this theme because the idea has to be implemented. At a higher level of innovation is creativity where the teacher has the capability of putting information and experience together into new configurations.
**Gestalt** – The Gestalt theme indicates the teacher has a drive toward completeness. The teacher sees in patterns and is uneasy until work is finished. When Gestalt is high, the teacher tends toward perfectionism. Even though form and structure are important, the individual student is considered first. The teacher works from individual to structure.

**Objectivity** – Objectivity is indicated when a teacher responds to the total situation. This teacher gets facts and understands first as compared to making an impulsive reaction.

**Focus** – Focus is indicated when a person has models and goals. The person’s life is moving in a planned direction. The teacher knows what the goals are and selects activities in terms of these goals.
APPENDIX B

SAMPLE TEACHER PERCEIVER INTERVIEW QUESTIONS AND SAMPLE RESPONSES
To conceptualize the format of the Teacher Perceiver Interview, a sample framework is provided depicting one hypothetical question and one hypothetical response for each of the twelve Teacher Perceiver Themes. Each hypothetical question might be similar to, but is not a reproduction of, the actual Teacher Perceiver Interview. The responses are merely paraphrased definitions of each theme. Definitions of each theme are found in Appendix A.

1. Theme: Mission
   
   Question: What has motivated you to choose teaching as a profession?
   
   Response: I believe that all students can grow and learn.

2. Theme: Empathy
   
   Question: How would you handle a student saying, “This assignment is stupid”?
   
   Response: I would try to see through the eyes of the student.

3. Theme: Rapport Drive
   
   Question: What kind of relationship would you like to have with your students?
   
   Response: I would like an approving and mutually favorable relationship with my students.
4. Theme: Individualized Perception

   Question: How do you meet the different needs of each of your students?
   
   Response: I will personalize each student’s educational program to meet his/her specific needs.

5. Theme: Listening

   Question: How do you know you are a good listener?
   
   Response: I listen to others with responsiveness and acceptance.

6. Theme: Investment

   Question: What do you want to receive most as a teacher?
   
   Response: I want to feel satisfaction in seeing my students grow.

7. Theme: Input Drive

   Question: Where do you get your ideas when helping students learn?
   
   Response: I am always seeking materials and relevant experiences to help my students learn.

8. Theme: Activation

   Question: How do you assist students achieve desired goals?
   
   Response: I stimulate students to think, respond, feel, and learn.
9. Theme: Innovation

Question: Explain an activity that has involved taking a different approach?

Response: I use specific new ideas and techniques that put information and experience together into new configurations.

10. Theme: Gestalt

Question: What process do you use to prepare for class?

Response: I organize myself to best suit students first. I tend toward perfectionism.

11. Theme: Objectivity

Question: How would you handle a situation where two students were fighting?

Response: I would first gather the facts and then take time to process the information rather than making an impulsive reaction.

12. Theme: Focus

Question: What are some of your future professional goals?

Response: I have goals and select various activities that will allow me to achieve my goals.
APPENDIX C

SAMPLE TEACHER RATING SCALE

STUDY I
TEACHER RATING SCALE

Teacher ______________________

Directions:

This survey contains 10 questions. Please read each of the following questions carefully. For each question, circle ONE numeric rating that best describes the teacher being evaluated. Please return this form directly to __________, Director of Personnel.

Mission

To what extent does the teacher believe that students seek to achieve all that they are capable of achieving?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Investment

To what extent does the teacher receive satisfaction in seeing students learn and grow?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

155
**Objectivity**

To what extent does the teacher first obtain facts before making a decision?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

**Empathy**

To what extent does the teacher attempt to see situations through the eyes of students?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

**Rapport Drive**

To what extent does the teacher seek to maintain mutually favorable relationships with students?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

**Activation**

To what extent does the teacher stimulate students to think and learn?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Innovation

To what extent does the teacher use new ideas that combine information and experience for students to learn?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Gestalt

To what extent does the teacher place students first when preparing for class?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Input Drive

To what extent does the teacher seek materials and relevant experiences to help students learn?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Individualized Perception

To what extent does the teacher individualize each student’s educational program to meet specific needs?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

157
APPENDIX D

SAMPLE TEACHER RATING SCALE
STUDY II
TEACHER RATING SCALE

Teacher ____________________

Directions:

This survey contains 12 questions, one question for each of the Teacher Perceiver Themes. Please read each of the following questions carefully. For each question, circle ONE numeric rating that best describes the teacher being evaluated. Please return this form directly to ____________, Coordinator of Human Resources.

Mission

To what extent does the teacher believe that students seek to achieve all that they are capable of achieving?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Investment

To what extent does the teacher receive satisfaction in seeing students learn and grow?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Focus

To what extent does the teacher strive toward achieving professional goals?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Objectivity

To what extent does the teacher first obtain facts before making a decision?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Listening

To what extent does the teacher listen to others with responsiveness and acceptance?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Empathy

To what extent does the teacher attempt to see situations through the eyes of students?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
**Rapport Drive**

To what extent does the teacher seek to maintain mutually favorable relationships with students?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

**Activation**

To what extent does the teacher stimulate students to think and learn?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

**Innovation**

To what extent does the teacher use new ideas that combine information and experience for students to learn?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

**Gestalt**

To what extent does the teacher place students first when preparing for class?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
**Input Drive**

To what extent does the teacher seek materials and relevant experiences to help students learn?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**Individualized Perception**

To what extent does the teacher individualize each student’s educational program to meet specific needs?

<table>
<thead>
<tr>
<th>Low Evidence</th>
<th>Moderate Evidence</th>
<th>High Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
References


Myart v. Motorola, 110 Congressional Record 5662-64 (1964).


