

3

92

18814

U·M·I

MICROFILMED 1992

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

PREVIEW

Order Number 9218814

**An evaluation of the admissions testing program of the Saint
Thomas Choir School, Manhattan**

Kraus, Susan, Psy.D.

Pace University, 1993

U·M·I

**300 N. Zeeb Rd.
Ann Arbor, MI 48106**

PREVIEW

**AN EVALUATION OF THE ADMISSIONS TESTING PROGRAM
OF THE SAINT THOMAS CHOIR SCHOOL, MANHATTAN**

by

Susan Kraus

**A Doctoral Project Submitted in Partial Fulfillment of
the Requirements for the Degree of Doctor of Psychology
in the Department of Psychology at Pace University**

NEW YORK

1993



PSYCHOLOGY DEPARTMENT
PSY.D. PROJECT
FINAL APPROVAL FORM

(Please type all information)

NAME: Susan Kraus

TITLE OF PROJECT: An Evaluation of the Admissions Testing
Program of the Saint Thomas Choir School,
Manhattan

DOCTORAL PROJECT COMMITTEE:

PROJECT ADVISOR: Jack L. Herman, Ph.D.
(Name)
Professor, Pace University
(Title) (Affiliation)

PROJECT CONSULTANT: Alfred W. Ward, Ph.D.
(Name)
Associate Professor, Pace University
(Title) (Affiliation)

FINAL APPROVAL OF COMPLETED PROJECT:

I have read the final version of the doctoral project and certify that it meets the relevant requirements for the Psy.D. degree in School-Community Psychology.

Jack L. Herman
(Project Advisor's Signature)

3/16/92
(Date)

Alfred W. Ward
(Project Consultant's Signature)

3/16/92
(Date)

Acknowledgements

I would like to express my special gratitude to five people who helped and encouraged me throughout the process of completing this doctoral project: my Advisor, Professor Jack L. Herman; my Consultant, Professor Alfred W. Ward; Gordon H. Clem, Headmaster of the Saint Thomas Choir School; Joseph L. Casalvieri, Computer Specialist at the Saint Thomas Choir School; and my husband, Donald C. Kraus, who tirelessly shared his editorial skills with me.

TABLE OF CONTENTS

Chapter	Page
Title Page	i
Approval Page	ii
Acknowledgments	iii
Table of Contents	iv-v
List of Tables	vi
Abstract	vii-viii
 I. Introduction	 1
General Introduction	
Review of the Literature	
Statement of Purpose	
II. Methodology	36
Subjects	
Materials	
Procedure	
III. Results	48
Description of Sample	
Evaluation of Research Questions	
Additional Analyses	

IV. Discussion	66
Research Questions	
Implications and Practical Suggestions	
Suggestions for Future Research	
References	74
Appendices	78
A. The Developmental Bender Scoring System	
B. Emotional Indicators on the Bender Test for Children	
C. Koppitz Emotional Indicators on HFDs of Children	
D. Rank-Ordered List of Choir Ratings Assigned to Sample	
E. Five Categories of Choir Ratings and Their Composition	
F. Subject Grades Assigned to Students	
G. Behavior Checklist	
H. Ratings by Faculty and Housemothers of Social/Emotional Outcome Variables	
I. Pilot Study: Semi-Random Choice of Student's Advisor Letters	
J. Data Sheet for Pilot Study - Interrater Reliability	
K. Percentage of "Not Mentioned" Ratings - Pilot Study	
L. Letter of Permission to Conduct Research	
M. Descriptive Statistics for Predictor Variables	
N. Descriptive Statistics for Outcome Variables	

LIST OF TABLES

Table		Page
Table 1.	Correlations among Predictor and Outcome Variables for Academic Year 1	51
Table 2.	Correlations among Predictor and Outcome Variables for Academic Year 2	52
Table 3.	Correlations among Predictor and Outcome Variables for Academic Year 3	53
Table 4.	Correlations among Predictor and Outcome Variables for Academic Year 4	54
Table 5.	Multiple Regression Analyses: Academic Predictors and Academic Outcome Variables	56
Table 6.	Multiple Regression Analyses: Academic Predictors and Social/Emotional Outcome Variables	57
Table 7.	Multiple Regression Analyses: Social/Emotional Predictors and Academic Outcome Variables	58
Table 8.	Multiple Regression Analyses: Social/Emotional Predictors and Social/Emotional Outcome Variables	59

ABSTRACT

The Saint Thomas Choir School is an independent boarding school for boys in grades 5 through 8 who sing in the boy choir of Saint Thomas Church, Manhattan. After passing a vocal audition, a candidate for admission is seen by a school psychologist, who administers the WISC-R, the Bender, and the House-Tree-Person drawings. The Headmaster uses these test results in making admissions decisions.

This study evaluated how well student behavior in several key areas was predicted by information obtained from the admissions tests. Test results included WISC-R Full Scale, Verbal, and Performance IQ scores; Bender Developmental scores (Koppitz); Bender Emotional Indicator scores (Koppitz); Emotional Indicator scores (Koppitz) and Naglieri Draw A Person scores on the male Human Figure Drawings. Student behavior was measured in several ways: final yearly grades; need for individual academic help from a learning specialist; need for individual psychotherapy; willingness to do work; and yearly choir ratings.

Archival data were used, with anonymity guaranteed, on 87 students enrolled in the graduating classes of 1982 through 1990. Straight correlations and a series of multiple regression analyses were performed to determine how

well outcome variables were predicted by test results, over four years. Groups receiving and not receiving the services of individual academic help and psychotherapy were also examined and compared with each other on key variables.

Results indicated that grades in academic subjects for all four years were very well predicted ($p < .001$ level) by WISC-R Full Scale IQ scores. In some cases grades were also predicted at a significant level by Bender Developmental scores (Koppitz). Need for individual academic help was negatively correlated with WISC-R Full Scale IQ scores. Need for individual psychotherapy was not correlated with any admissions test data. Willingness to do work was not predicted by any admissions test data, except in year 4 only by the Naglieri Draw A Person scores. Success in choir was not predicted by any admissions test data, except in year 4 only by the Naglieri Draw A Person scores and the Bender Developmental scores (Koppitz).

The conclusions of this research were the following recommendations to the Saint Thomas Choir School: 1) to continue to use the WISC-R (or the recent WISC-III) as an admissions test, as it has predicted academic success very well in past years; and 2) to add more sensitive instruments for personality assessment to the current test battery in order to better predict student behavior.

Chapter I

General Introduction

The research described below is an evaluation, using archival data, of the admissions testing program at the Saint Thomas Choir School in Manhattan. This research has practical implications for this particular school and its admissions testing program. Also, because of the unique nature of the school, this project has provided information not usually available on the predictive validity of three commonly used assessment instruments: the Bender Visual-Motor Gestalt Test, the Wechsler Intelligence Scale for Children - Revised, and the Human Figure Drawing.

The Saint Thomas Choir School is a boarding school for boys in grades 5 through 8, who sing in the boy choir of Saint Thomas Episcopal Church in Manhattan. Singing is a major part of the boys' curriculum, and vocal aptitude is a necessary requirement for admission to the school. The students rehearse ten hours each week and sing six weekly church services, in addition to three to four annual concerts, yearly tours, recordings, and television appearances.

The Choir School is a fully accredited independent school with a curriculum including English, mathematics, history, science, art, Latin, French, music theory, and

musical instruments. There are six full-time and ten part-time instructors. As is the case in many independent schools, the faculty are not New York State Education Department licensed teachers, but are graduates of bachelors and masters degree college and university programs in the fields which they teach.

The maximum student body is fifty; currently there are forty-five students in the four grades. Recent graduating classes range in size from four to fifteen. Each student receives a great deal of individualized instruction and is able to progress in each subject at his own rate. All of the students are known very well and are regularly evaluated on several behavioral dimensions.

All Choir School students are boarding students. There is an emphasis placed on developing a close, caring community among the boys and the seven faculty and three housemothers who live in the school building. Each boy has a faculty advisor and writes a weekly journal entry about his feelings and experiences to which his advisor responds in writing and in personal discussion.

The admissions process begins with a vocal audition, usually in the 4th grade year, conducted by the Choir Master. If the applicant demonstrates adequate musical and vocal aptitude, he is then scheduled for individual psychological assessment by a school psychologist. If he does not pass his vocal audition, he is not admitted.

The admissions testing battery consists of the Bender Visual-Motor Gestalt Test, the WISC-R, and the House-Tree-Person drawings. If test results are adequate, the boy spends a "trial week" at the school, fully participating in all activities and living at the school, helped by an assigned student guide. A decision is then made by the Headmaster and faculty about the candidate's admission.

The psychological assessment battery is relied on heavily in screening candidates for admission. The Bender is used as both a neuropsychological screening test and as a projective test. The school psychologist notes features of the reproduced designs, such as distortion, poor organization, and very large or very small designs, which may suggest a possible learning disability and/or emotional disturbance.

If a boy's WISC-R suggests that he lacks sufficient aptitude for the rigorous academic program or has serious learning problems, he will most likely not be admitted. There are no absolute cut-off scores in this regard, and prior academic record is also considered when WISC-R scores are too low. However, a Full Scale WISC-R score below 120 may very well lead to a candidate's rejection. A large difference between a student's Verbal and Performance IQ scores and very low subtest scores are also noted.

The House-Tree-Person drawings are used as screening instruments regarding each boy's emotional maturity and well-being. Admission to the Choir School means leaving home at the age of ten and adjusting to life in a community with more than forty boys and with many adults. A boy whose drawings suggest emotional difficulties may be denied admission or may be accepted under the condition that he be given a thorough psychological evaluation, followed by weekly psychotherapy, if recommended. To date, drawings have not been scored, but have been inspected and globally assessed through the clinical judgment of the examiner.

The unique nature of the Saint Thomas Choir School has provided the opportunity for testing how well commonly used psychological assessment instruments have predicted behavior of several types over four years. This school is a rare environment, and though this limits the generalizability of findings to other schools, its special features make it similar to a real-life laboratory setting in which a longitudinal study has taken place.

First of all, in order to enter the "experiment," each student has taken an abbreviated psychological test battery. It is typically only students who have been referred to a psychologist because of academic or emotional problems who receive individually administered psychological tests. Thus, there are more test data available on "normal" children at this school than are typically available in other schools.

Secondly, there are unusually detailed records of student behavior at the Choir School. This is true in part because of the high teacher:student ratio. Also, because all of the students are boarding students and because the full-time faculty live at the school, many of the staff members who have evaluated the students have lived with them twenty-four hours a day. They have had the opportunity to see and interact with the students in many activities. They have known each student better than a teacher of day students would be likely to know his/her students.

Finally, a very high proportion of the evaluators in this study have been with the Choir School for many years. The Choir Master has been at the school for twenty years. The school psychologist who does the admissions testing has worked at the school for nearly thirty years. The Headmaster has been on the faculty for over thirty years and has been Headmaster for more than twenty years. Three of the other five full-time faculty members have been with the school for more than fifteen years. Only two major academic subjects, English and science, have been taught by more than one teacher during the years of the present study, and two learning specialists have been employed during those years. Such continuity and constancy are rare.

Review of the Literature

School psychologists frequently use abbreviated psychological test batteries in the evaluation of school-age

children. These typically consist of a perceptual test, an individually administered intelligence test, a personality test, and an academic achievement test. The results from such test batteries are used most often for purposes of educational planning and placement in the schools. More extensive testing is certainly necessary for the purpose of diagnosing serious neurological, cognitive, or emotional disorders and disturbances.

It is assumed that test results provide adequate information upon which decisions about school placement can be made. This entails assumptions about the tests' reliability, validity, and usefulness in predicting behavior, or criterion-related validity. A discussion follows of general issues regarding criterion-related validity, as this study investigated the predictive validity of three commonly used psychological tests. This discussion will be followed by a review of the relevant literature on the tests used by the Saint Thomas Choir School for admissions purposes: the Bender Visual-Motor Gestalt Test, the Wechsler Intelligence Scale for Children - Revised (WISC-R), and the House-Tree-Person drawings (H-T-P).

Criterion-related validity is concerned with how effective scores on a test are in predicting behavior in specific areas. Test performance is compared with a criterion, or independent measure of the behavior which the test is expected to predict. Concurrent validity is

examined when the criterion used for validation is measured at the same time the test is administered. Predictive validity is examined when the criterion used for validation is measured at one or more intervals of time after the test is administered.

The criteria used for test validation may be relatively objective measures of behavior. For example, intelligence tests are often validated using measures of academic achievement, such as scores on academic achievement tests and grades in school. Ratings which involve the relatively subjective judgment of the raters, such as teacher ratings of student personality characteristics, are also often used as validation criteria. Ratings "are particularly useful in providing criteria for personality tests, since objective criteria are ... difficult to find in this area" (Anastasi, 1988, p. 151).

"Criterion-related validity is often used in local validation studies, in which the effectiveness of a test for a specific program is to be assessed. ... Criterion-related validity can be best characterized as the practical validity of a test for a specified purpose" (Anastasi, 1988, p. 151). In the case of tests used to select students for admission to schools, predictive validity is most important.

The Bender Visual-Motor Gestalt Test: A Neuropsychological Test

The Bender Visual-Motor Gestalt Test consists of "nine designs originally used to demonstrate the tendency of the perceptual system to organize visual stimuli into Gestalten (configurational wholes)" (Lezak, 1983, p. 385).

Dr. Lauretta Bender assembled these designs (1938) and used them to study the mental development of children. The designs have been reproduced on cards which are presented one at a time for the subject to copy, with the designs in view, on a blank sheet of eight and a half inch by eleven inch paper. The Bender, "one of the most popular and widely administered psychological tests," (Rossini & Kaspar, 1987, p. 254) is given to people of all ages, and has been used primarily to provide information regarding subjects' developmental delays, neurophysiological impairment, mental retardation, and emotional disorders (Siegel, 1987; Koppitz, 1963).

Successful reproduction of the Bender designs requires that the subject see the stimulus cards, understand his/her perceptions, and then translate these perceptions into controlled, coordinated motor actions. Thus, the Bender requires both receptive and expressive capabilities (Siegel, 1987). These abilities develop over time, and expectations of performance on the Bender thus change with age.

Dana, Field, and Bolton (1983) reviewed the literature on variations in the administration, scoring, and applications to particular populations of the Bender. They stated, regarding scoring, that "some clinicians prefer intuitive and global approaches while others advocate objective systems" (p. 78). The debate as to which approach yields the fullest, most reliable information continues. Objective scoring systems, however, are used in most of the published research studies on the Bender's reliability and validity.

Koppitz (1963) standardized the Bender for use with children. 1,104 public school children ranging in age from 5 years to 10 years 11 months, in kindergarten to fifth grade, were tested individually. These children attended schools located in rural, small town, suburban, and urban areas of midwestern and eastern states. Their Bender protocols were scored by Koppitz using her Developmental Bender Scoring System (Koppitz, 1963, pp. 16-32; see Appendix A), a scoring system which continues to be widely accepted (Dana, Field, & Bolton, 1983, p. 80).

Koppitz computed mean composite scores and standard deviations for boys and girls for each six month age group in her normative sample (Koppitz, 1963, p. 188). She reported that scores for boys and girls did not differ to a statistically significant extent at any age level (Koppitz, 1963, p. 35). Thus, one set of scores per age level is

given for all children. By referring to these normative data, the examiner can determine whether a child's visual-motor perception is within normal limits for his/her age group.

Although Koppitz attempted to include children from various areas and socioeconomic groups in her 1963 normative sample, this sample in fact included very few minority children. These normative data were therefore typical of only white American school children. In 1974 Koppitz conducted a new normative study of 975 elementary school students, ranging in age from 5 to 11, from the West, South, and Northeast. These children, like those in the earlier study, lived in rural areas, small towns, suburbs, and metropolitan centers. The 1974 sample was 86% white, 8.5% black, 4.5% Mexican-American and Puerto Rican, and 1% Asian. Koppitz computed mean composite scores and standard deviations for children in each six month age group of this sample (1975, p. 185), and these data may be used as the earlier data had been in the assessment of children.

Reliability

Summarizing research on the Bender conducted between 1962 and 1980, Tolor and Brannigan (1980) report that "research on reliability without exception demonstrates that satisfactory inter-rater and test-retest reliabilities are obtained with objective scoring systems and for raters having varying degrees of clinical sophistication. However,

the research also suggests that one should rely on 'total' errors rather than 'specific' items or error types when interpreting an individual's protocol" (p. 178).

Koppitz (1975) reported the results of 23 studies conducted between 1965 and 1973 of the interscorer reliability of her Developmental Bender Scoring System. Thirty-one interscorer correlations are presented, ranging from .79 to .99, with 81% at .89 or higher (p. 19). Koppitz (1975) also reported on nine research studies regarding test-retest reliability of Developmental Bender Scores. She stated that "research findings showed conclusively that Bender Test scores of normal elementary-school pupils (end of kindergarten to sixth grade) are reliable" (p. 30).

Validity

Considerable research has been done to assess the Bender's validity for many uses. Dana, Field, and Bolton (1983) reviewed the literature and concluded that "there is consensus that the [Bender] may be used to diagnose organic brain pathology or neurological impairment" (p. 81). They also state that "responsible clinicians concur in finding ... intelligence, personality dynamics, psychopathology, and developmental process as legitimate assessment outcomes," though there is less agreement concerning these diagnoses than concerning organicity (p. 81). Research has also been done on the relationship between Bender scores and academic achievement.