DISCRIMINANT VALIDITY OF THE DRAW A PERSON SCREENING PROCEDURE FOR EMOTIONAL DISTURBANCE FOR INCARCERATED JUVENILE DELINQUENTS IN SPECIAL EDUCATION

DISSERTATION

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

By

Mark Alex Matavich, M.A.

*****

The Ohio State University
1998

Dissertation Committee:
Jack A. Naglieri, Ph.D., Adviser
Timothy E. Heron, Ed.D.
Joe E. Wheaton, Ph.D.

Approved by

Adviser
College of Education
ABSTRACT

This study examined the effectiveness of the Draw-A-Person Screening Procedure for Emotional Disturbance (DAP: SPED) differentiating juvenile delinquents with and without a history of sex offenses from normal juveniles. The DAP: SPED was administered to 78 juvenile delinquents with sex offense, 76 juvenile delinquents without sex offense, and 154 normal adolescents (ages 13-17) matched on age, race, and gender. The normal or control group was derived from the DAP: SPED national standardization sample.

Analyses of the results revealed overall group differences between juvenile delinquents with and without a history of sex offenses and controls (normals) on the 59-item DAP: SPED Scale. The DAP: SPED mean T-score earned by the 154 juvenile delinquents (M = 58.11, SD = 7.99) was significantly higher (F = 197.22, p < .0001) than the mean T-score earned by the 154 controls or normal adolescents (M = 46.25, SD = 6.78). The DAP: SPED mean T-score earned by the 78 juvenile delinquents with sex offense (M = 58.08, SD = 7.77) was significantly higher (F = 142.48, p < .0001) than the mean T-score earned by the 154 controls. Also, the DAP: SPED mean T-score earned by the 76 juvenile delinquents without sex offense (M = 58.14, SD = 8.26) was significantly higher (F = 61.14, p < .0001) than the mean T-score earned by the 154
controls. No significant difference ($F = .00, p < .9582$) was noted between the juvenile delinquents with sex offense ($M = 58.08$, $SD = 7.77$) and juvenile delinquents without sex offense ($M = 58.15$, $SD = 8.26$) on this scale.

The juvenile delinquents with sex offense and juvenile delinquents without sex offense scored higher on the Aggression Scale than the controls. Interestingly, the juvenile delinquents without sex offense scored higher on this scale than juvenile delinquents with sex offense. The Sex Scale used in this study was not able to differentiate between any of the groups.

Results of sensitivity and specificity analysis (true positives and true negatives) suggest that identification of group membership was most accurate using the Total DAP: SPED cut-off score of 55 where 55% of the juvenile delinquents and 93% of the controls were correctly identified. Finally, all group comparisons with significant $F$-test results yielded large mean score effect size differences.

These findings suggest that the DAP: SPED demonstrates some utility and validity differentiating juvenile delinquents with and without a history of sex offenses from normal juveniles or adolescents.
This study is dedicated to my parents, Alex and Jane Matavich.
ACKNOWLEDGEMENTS

I would like to extend my thanks to the Ohio Department of Youth Services, namely Mr. Gerald Plassenthal, Mr. Geno Natalucci-Perischetti, Dr. Lee Underwood, and Dr. Alice P. Franklin-Elder for their support of this research project.

A sincere thank you and appreciation is extended to my supportive committee led by Dr. Jack A. Naglieri, adviser, Dr. Timothy E. Heron, and Dr. Joe E. Wheaton.

Many, many thanks to the technical/professional support staff of Linda Armstrong, Dennis Koenig, and Charlotte Danniger.

Finally, I am grateful to my family and very special circle of friends who served as constant reminders of the benefits of the road less traveled.
VITA

Academic Background

1985................. M. A. The University of Akron
Graduate Research Associate
Akron, Ohio

1980................. B. A. Youngstown State University
Major: Psychology
Minors: Economics and Management
Youngstown, Ohio

Professional Experience

1992-present.......... School Psychologist, Ohio Department of Youth Services,
Delaware, Ohio

1987-present.......... School Psychologist, Columbus Public Schools,
Columbus, Ohio

1986-87............... School Psychologist, Charlotte-Mecklenberg County Public
Schools, Charlotte, North Carolina

1985-86............... Intern School Psychologist, Canton City Schools, Canton,
Ohio

Publications

Diagnostique, 17(3), 195-201.

performance on the WISC-III and the Stanford-Binet Intelligence Scale: Fourth
FIELDS OF STUDY

Major Field: School Psychology
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>Historical Perspective</td>
<td>1</td>
</tr>
<tr>
<td>Issues and Controversies</td>
<td>4</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>6</td>
</tr>
<tr>
<td>Importance of Identifying Sex Offenders</td>
<td>6</td>
</tr>
<tr>
<td>Research Questions</td>
<td>7</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>10</td>
</tr>
<tr>
<td>2. REVIEW OF RELATED LITERATURE</td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>12</td>
</tr>
<tr>
<td>Human Figure Drawings as a Projective Tool</td>
<td>13</td>
</tr>
<tr>
<td>Machover's Work</td>
<td>13</td>
</tr>
<tr>
<td>Koppitz's Work</td>
<td>16</td>
</tr>
<tr>
<td>Schildkrout, Shenker, and Sonnenblick's Work</td>
<td>19</td>
</tr>
<tr>
<td>Naglieri, McNeish, and Bardos' Work</td>
<td>19</td>
</tr>
<tr>
<td>Item Level and Global Interpretation</td>
<td>21</td>
</tr>
<tr>
<td>Human Figure Drawings and Sex Offenders</td>
<td>22</td>
</tr>
</tbody>
</table>

viii
5. DISCUSSION

Group Differences on the Total DAP-SPED .................................................. 60
Group Differences on the Sex Scale............................................................. 62
Group Differences on the Aggression Scale.................................................. 63
Sensitivity and Specificity Analysis .............................................................. 63
Conclusion ...................................................................................................... 65
Limitations of Study ....................................................................................... 66
Suggestions for Future Research ................................................................. 67

LIST OF REFERENCES ..................................................................................... 68

APPENDIX A  Sex Scale .................................................................................... 76
APPENDIX B  Aggression Scale ....................................................................... 78
APPENDIX C  Examples of Human Figure Drawings ...................................... 80
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Demographics of Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense</td>
<td>33</td>
</tr>
<tr>
<td>2. Geographic Regions where Controls (Standardization Sample) and Juvenile Delinquents were Drawn</td>
<td>34</td>
</tr>
<tr>
<td>3. Mean and Standard Deviation Scores for Controls and all Juvenile Delinquents on the Total DAP:SPED Raw Score Scale and the Total DAP:SPED T-Score Scale</td>
<td>40</td>
</tr>
<tr>
<td>4. Mean and Standard Deviation Scores for Controls and all Juvenile Delinquents on the Sex and Aggression Scales</td>
<td>41</td>
</tr>
<tr>
<td>5. Raw Score and T-Score Means with Standard Deviations on the Total DAP:SPED for Controls and Juvenile Delinquents with Sex Offense</td>
<td>45</td>
</tr>
<tr>
<td>6. Mean and Standard Deviation Scores for Controls and Juvenile Delinquents with Sex Offense on the Sex and Aggression Scales</td>
<td>46</td>
</tr>
<tr>
<td>7. Raw Score and T-Score Means with Standard Deviations for Controls and Juvenile Delinquents without Sex Offense</td>
<td>50</td>
</tr>
<tr>
<td>8. Mean and Standard Deviation Scores for Controls and Juvenile Delinquents without Sex Offense on the Sex and Aggression Scales</td>
<td>51</td>
</tr>
<tr>
<td>9. Raw Score and T-Score Means with Standard Deviations on the Total DAP:SPED for Juvenile Delinquentd with Sex Offense and Juvenile Delinquentd without Sex Offense</td>
<td>55</td>
</tr>
<tr>
<td>10. Mean and Standard Deviation Scores for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense on the Sex and Aggression Scales</td>
<td>56</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Controls and all Juvenile Delinquents</td>
</tr>
<tr>
<td>2.</td>
<td>Sensitivity and Specificity Analysis of the Sex Scale for Controls and all Juvenile Delinquents</td>
</tr>
<tr>
<td>3.</td>
<td>Sensitivity and Specificity Analysis of the Aggression Scale for Controls and all Juvenile Delinquents</td>
</tr>
<tr>
<td>4.</td>
<td>Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Controls and Juvenile Delinquents with Sex Offense</td>
</tr>
<tr>
<td>5.</td>
<td>Sensitivity and Specificity Analysis of the Sex Scale for Controls and Juvenile Delinquents with Sex Offense</td>
</tr>
<tr>
<td>6.</td>
<td>Sensitivity and Specificity Analysis of the Aggression Scale for Controls and Juvenile Delinquents with Sex Offense</td>
</tr>
<tr>
<td>7.</td>
<td>Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Controls and Juvenile Delinquents without Sex Offense</td>
</tr>
<tr>
<td>8.</td>
<td>Sensitivity and Specificity Analysis of the Sex Scale for Controls and Juvenile Delinquents without Sex Offense</td>
</tr>
<tr>
<td>9.</td>
<td>Sensitivity and Specificity Analysis of the Aggression Scale for the Controls and Juvenile Delinquents without Sex Offense</td>
</tr>
<tr>
<td>10.</td>
<td>Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>11. Sensitivity and Specificity Analysis of the Sex Scale for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense</td>
<td>58</td>
</tr>
<tr>
<td>12. Sensitivity and Specificity Analysis of the Aggression Scale for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense</td>
<td>59</td>
</tr>
<tr>
<td>13. Illustrates emotional indicators present in a human figure drawing of a juvenile delinquent with sex offense</td>
<td>82</td>
</tr>
<tr>
<td>14. Illustrates emotional indicators present in a human figure drawing of a juvenile delinquent without sex offense</td>
<td>84</td>
</tr>
<tr>
<td>15. Illustrates the human figure drawing of an adolescent from the control group</td>
<td>86</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Historical Perspective

Human figure drawings have been used to assess children and adolescents’ development and personality characteristics since the late nineteenth century. In 1885, Ebenezer Cooke suggested that children’s drawings show successive stages of development. Shortly thereafter Corrado Ricci’s (1887) work on Italian children’s drawings pioneered the scientific interest in human figure drawings for the twentieth century (cited in Harris, 1963). These and other studies demonstrated that the nature and content of drawings were dependent primarily upon cognitive development.

Lamprecht (1907) attempted to determine if a relationship existed between aptitude in drawing and general intellectual ability as indicated by schoolwork. Ivanhoff (1909) followed and devised a method of scoring the drawings that resulted in positive correlations between the obtained values and teacher ratings of general ability (as cited in Harris, 1963). When studying aspects of drawings related to intellectual ability, other researchers described changes in children’s drawings that reflected emotional rather than intellectual factors (Luquet, 1927, 1929).
In her classic work, *Measurement of Intelligence by Drawings*, Goodenough (1926) credits Barnes (1892) as the first to analyze and classify children based on details of their drawings. Goodenough devised a quantitative approach and articulated a process to describe the Draw-A-Man. This scoring system for estimating individuals’ levels of intellectual functioning was widely used for many years. Goodenough (1926) discovered that differences in drawings appeared unrelated to development or intellectual levels which led her to state that the human figure drawing contains signs that “probably carry profound meaning, had we but the wisdom to understand it” (p. 60).

An approach to uncover the meaning of HFD was provided in Machover’s (1949) publication *Personality Projection in the Drawing of the Human Figure*. Her approach has been a popular method for personality assessment (Cummings, 1986), ranging in position from 2nd to 11th among the most frequently used psychological instruments between 1959 and 1984 (Brown & McGuire, 1976; Wade & Baker, 1977) and ranking 8th in the most recent survey conducted by Lubin, Larsen, and Matarazzo (1984).

Machover suggested that human figure drawings include signs of specific interpersonal problems and expressions of individual mood. Based on her clinical experience with emotionally disturbed male adolescents and adult institutionalized psychotics, Machover focused on the projective aspects of drawings stating, “the human figure drawn by an individual who is directed to ‘draw a person’ relates intimately to the impulses, anxieties, conflicts, and compensations characteristic of that individual” (Machover, 1949, p. 35). The essential focus of her technique was the interpretation and
comparison of individual signs and various configurations of these signs based on clinical interpretation rather than normative data.

Koppitz (1968, 1984) developed a system for interpreting emotional signs in human figure drawing that provided detailed descriptions for each item. Her aim was to identify both the signs that occurred infrequently in normal children as a guide for considering atypical performance and by using the total number of items present. The total number of items present per drawing was considered more important than the interpretation of individual items in Koppitz’s scoring system. This is in contrast to those who used the presence of individual items as signs of emotional disturbance (Machover, 1948; Hammer, 1958; Urban, 1963).

Item level analysis, or interpretation of drawings such as small size suggesting depression or a slash line for a mouth associated with verbal aggression or sadistic personality (Machover, 1949), has been the subject of many research studies over the years. Kahill (1984) found conflicting evidence for the one-to-one relationships between aggression and specific body parts (e.g., teeth, mouth, arms, hands, and fingers). Montague and Prytula (1975) disputed Machover’s hypothesis that delinquents drew, for example, figures with hands in pockets. They found that this item did not differentiate between delinquent and nondelinquent groups.

Griffin and Lemley (1967) emphasized a cluster of signs be used for identification of emotional problems although others found differences based on individual signs (Daum, 1983). While there may not be a direct relationship between an emotional sign and a specific behavior, Saarni and Azara (1977) suggested a cluster of items indicated
conflict about hostility and aggression. These studies (Roback, 1967; Kahill, 1984) suggest, and are supported by other researchers (Tharington & Stark, 1990; Naglieri, McNeish, and Bardos, 1991; Naglieri & Pfeiffer, 1992) that a global approach for interpreting emotional signs in human figure drawings could be effective.

Recently, a holistic approach to interpretation of human figure drawings has been provided (Naglieri, McNeish, & Bardos, 1991) that represents a shift from high-level inference about individual signs interpretation and abnormality (Haddad, 1995). According to these authors, it is the accumulation of a number of signs and not individual signs on the human figure drawing that may provide a more accurate understanding of a child or adolescent’s overall level of emotional functioning. Their scoring system was designed to increase reliability, reduce the number of ambiguous items, and use an actuarial approach to normative score generation.

Issues and Controversies

An exhaustive review of studies of human figure drawings between 1945 and 1956 conducted by Swensen (1957) revealed little support for Machover’s (1949) hypothesis that one projects specific intrapersonal or interpersonal conflicts onto drawings. A smaller section of studies was reviewed by Suinn and Oskamp (1969) where tenuous validity of human figure drawings as assessment instruments is reported. The empirical literature on the reliability and validity of human figure drawings between 1967 and 1982 was reviewed by Kahill (1984) who concluded that Machover’s original hypothesis is not supported by the content and structural aspects of drawings. Other
researchers (Dieffenbach, 1997; Mortensen, 1984; Phil & Nimrod, 1976), found no support for diagnostic validity of Koppitz's (1968) system of emotional indicators on drawings when subjected to empirical evaluation.

The literature addressing the use of human figure drawings as a projective personality assessment technique has been recently summarized by Naglieri, McNeish, and Bardos (1991) who said “....the research results have been characterized by conflict. Numerous case studies appear in the literature that attest to the technique's usefulness in isolated clinical situations, yet experimental research has offered little to no support for the technique's validity as an indicator of emotional or behavioral attributes” (p.2). Further, Roback (1968) stated “....perhaps in individual cases, they may provide insight into the drawer's perception of himself and others, but in these instances the cases are usually so extreme or the patient so disorganized that one could easily have gotten the same information from a multitude of other sources which would not have necessitated a testing situation” (pp. 16-17). According to Motta, Little, and Tobin (1993), “.....ample evidence exists that human figure drawings should not be used as personality test instruments in that they do not provide valid descriptions of personality, behavior, or social-emotional functioning” and question why human figure drawings continue to be among the most popular personality assessment devices (p. 165). This conclusion is in contrast to findings of other researchers who have found support for the global or holistic interpretation of emotional indicators in human figure drawings (Tharington & Stark, 1990; Naglieri, McNeish, and Bardos, 1991; Naglieri & Pfeiffer, 1992).
Significance of the Study

Even though Naglieri, McNeish, and Bardos (1991) suggest that differential diagnosis on the Draw-A-Person: Screening Procedure for Emotional Disturbance (DAP: SPED) may not be possible, this question has not been assessed using DAP:SPED items. Moreover, it is possible that the effectiveness of the DAP: SPED may be augmented with additional items. The intent of this study, therefore is to examine the effectiveness of selected items included in the DAP: SPED, additional items from the literature, and the total DAP: SPED scores for the identification of emotional problems of persons with juvenile delinquency with and without a history of sex offenses. This study will seek to determine whether some combination of sexual and aggressive items differentiates between juvenile delinquents with and without sex offenses. The items chosen are based on research evidence supporting them as sexual or aggressive in nature and their inclusion in DAP: SPED. Because of the paucity of research on juvenile sexual offenders, this study seeks to demonstrate that if a carefully selected sample of adolescents with specific problems are assessed with a good human figure drawing scoring system, then it is possible to screen for persons who have aggressive and atypical sexual tendencies. If such a system should show effectiveness then the identification of this important group could be facilitated.

Importance of Identifying Sex Offenders

Fingerhut and Kleinman (1990) report an unprecedented rise in rates of violent crimes among youth in the United States. Ever-increasing numbers of children and
adolescents serve as perpetrators of aggression, assault, and murder (Richters, 1993) and are depicted daily in newspapers, television news broadcasts, and rising statistics in the 1990's.

Clinical opinion rather than empirical research dominate the literature on juvenile sex offenders. Few studies have used comparison groups to identify characteristics which differentiate juvenile sex offenders from other adolescents, thereby reducing the opportunity to better identify and treat adolescents who commit sexual offenses. In fact, according to McGraw and Pegg-McNab (1989), “little has been reported on the psychological make-up of juvenile sex offenders” (p. 546). Further, “much of the research literature on adolescent sex offenders is relatively recent and much more remains to be done” (Davis & Leitenberg, 1987, p. 417) because “little is known about the characteristics of this population” (Kavoussi, Kaplan, & Becker, 1998, p. 241).

Research Questions

Controls vs. Juvenile Delinquents

1a. Is there a significant difference between controls and all juvenile delinquents on DAP:SPED Total T-Score?

1b. Is there a significant difference between controls and all juvenile delinquents on Sex Scale?

1c. Is there a significant difference between controls and all juvenile delinquents on Aggression Scale?
d. What is the sensitivity and specificity of the DAP:SPED Total T-Score, Sex Scale, and Aggression Scale using values of 0 through 3 SDs for controls and all juvenile delinquents?

e. What is the effect size difference between controls and all juvenile delinquents on all DAP:SPED items including additional sex and aggression items?

Controls vs. Juvenile Delinquents with Sex Offense

2a. Is there a significant difference between controls and juvenile delinquents with sex offense on DAP:SPED Total T-Score?

b. Is there a significant difference between controls and juvenile delinquents with sex offense on Sex Scale?

c. Is there a significant difference between controls and juvenile delinquents with sex offense on Aggression Scale?

d. What is the sensitivity and specificity of the DAP:SPED Total T-Score, Sex Scale, and Aggression Scale using values of 0 through 3 SDs for controls and juvenile delinquents with sex offense?

2e. What is the effect size difference between controls and juvenile delinquents with sex offense on all DAP:SPED items including additional sex and aggression items?

Controls vs. Juvenile Delinquents without Sex Offense

3a. Is there a significant difference between controls and juvenile delinquents without sex offense on DAP:SPED Total T-Score?
b. Is there a significant difference between controls and juvenile delinquents without sex offense on Sex Scale?

c. Is there a significant difference between controls and juvenile delinquents without sex offense on the Aggression Scale?

d. What is the sensitivity and specificity of the DAP:SPED Total T-Score, Sex Scale, and Aggression Scale using values of 0 through 3 SDs for controls and juvenile delinquents without sex offense?

e. What is the effect size difference between controls and juvenile delinquents without sex offense on all DAP:SPED items including additional sex and aggression items?

Juvenile Delinquents with Sex Offense vs. Juvenile Delinquents without Sex Offense

4a. Is there a significant difference between juvenile delinquents with sex offense and juvenile delinquents without sex offense on DAP:SPED Total T-Score?

b. Is there a significant difference between juvenile delinquents with sex offense and juvenile delinquents without sex offense on Sex Scale?

c. Is there a significant difference between juvenile delinquents with sex offense and juvenile delinquents without sex offense on Aggression Scale?

d. What is the sensitivity and specificity of the DAP:SPED Total T-Score, Sex Scale, and Aggression Scale using values of 0 through 3 SDs for juvenile delinquents with sex offense and juvenile delinquents without sex offense?
e. What is the effect size difference between juvenile delinquents with sex offense and juvenile delinquents without sex offense on all DAP:SPED items including additional sex and aggression items?

Definition of Terms

1. **Felony one offenses**: include murder, aggravated robbery (with weapon), and sale of certain drugs.

2. **Felony two offenses**: include robbery (no weapon), burglary, and felonious assault.

3. **Felony three offenses**: include possession of certain drugs, extortion, and grand theft auto.

4. **Felony four offenses**: include breaking and entering, vandalism, and grand theft.

5. **Sex Offenses**: include aggravated rape, sexual battery, gross sexual imposition, and felonious sexual penetration.

6. **Sex Scale**: (dependent variable) includes long nose, long feet, waistline emphasis, blank/hidden eyes, hidden/missing hands, crotch erasure, and crotch shading.

7. **Aggression Scale**: (dependent variable) includes slash mouth, large figure, aggressive symbols, talons, pointy foot/feet, teeth, and clenched fist.

8. **d-ratio**: an effect size statistic, which describes the difference between group means.
9. **Sensitivity Analysis**: proportion of cases correctly identified or true positives
   (accuracy of identification).

10. **Specificity Analysis**: proportion of non-cases correctly identified or true negatives (accuracy of identification).
CHAPTER 2

REVIEW OF RELATED LITERATURE

Background

Since the late nineteenth century, children's human figure drawings were used primarily as a tool to assess overall ability (Ricci, 1887; Sully, 1895). Barnes (1892) classified children by analyzing the details of their drawings following their depiction of the poem, "Johnny Look-in-the-Air" where meaning (projection) was derived from cloud configuration. Burt (1921) found drawings of children to be less dependent on learned skills and reported evidence of both developmental progress and delay when utilizing children's drawings to estimate ability.

Interest and investigation toward better comprehension of children's drawings continued to emerge in the 1990s. Seeking to use drawings primarily as a test of overall ability and mental maturity, practitioners followed Goodenough's (1926) method in the Draw-A-Man Test. Used in its original version for approximately 40 years, it was later revised and became the Goodenough-Harris Scoring System (Harris, 1963), restandardized (Harris & Roberts, 1972) and used as measures of children's intellectual ability. Subsequently, similar human figure drawing techniques have been offered as measures of intellectual development (Koppitz, 1968; McCarthy, 1972; Naglieri, 1988).
with evidence suggesting the techniques usefulness as a screening measure
of intellectual ability (Sattler, 1988).

Several researchers (Luquet, 1927, 1929; Lewis, 1928; Lembke, 1930 &
Lowenfield, 1947, 1952; Hammer, 1968) have recognized features in children's drawings
that suggest emotional status of the child (feelings, beliefs, and desires) in addition to
intellectual or developmental maturity.

**Human Figure Drawings as a Projective Tool**

**Machover's Work**

As a result of Goodenough's Draw-A-Man Test, interest in drawings as
personality assessment tools became popular in the 1940s. In fact, it was Machover
(1949) who first recognized children's anxieties, feelings, and conflicts were often
revealed in their drawings that were not necessarily related to intellectual ability.
Hammer (1968) noted qualitative characteristics such as boldness of line, size, position
on page, and facial expression seemed to offer pertinent clinical information about the
child where expressions of affect "flooded" (p. 154) the drawing. Machover, who was
influenced by psychoanalytic theory and personality analysis, was interested in the
pictorial representations of the drawer's impulses, conflicts, compensations, and feelings
in their drawings not necessarily related to mental ability. Her method was not a test and
had no formal scoring system and instead relied on the understanding, insight, and
experience of the clinician to interpret the social and contact features of human figure
drawings.
The drawing's head, for example, considered expressive and the center of social communication, purportedly revealed something about the child or adolescent. For example, an attempt to close out or remove oneself from the world was depicted by closed eyes, paranoia by eyes with excessive detail, and hostility by dark and emphasized eyes. Aggression was reportedly portrayed in a human figure drawing by a heavy single slash while an overly emphasized mouth suggested profanity, temper tantrums, or gastric symptoms. A child or adolescent's interactions with the environment are represented by the contact features such as hands, fingers, arms, feet, and toes. Lack of attention or omission of these items may be suggestive of a child's insecurity or problems dealing with sexual impulses.

Machover developed interpretive hypotheses for clothing as well suggesting, "irrelevant emphasis upon pockets is seen in drawings of infantile and dependent individuals" (1949, p. 79). Variables such as figure size, placement on the paper, and theme of the drawings were considered structural variables in her system. Focusing on the graphic features of drawings, Machover adhered to the belief that individual signs of drawings were directly related to specific emotional conflicts or disorders. The essential focus of her technique on the interpretation of individual signs and configurations was based on the comparison to theoretical constructs as opposed to normative data (McNeish, 1989).
It was Machover (1949) who offered the body image hypothesis as the most basic assumption underlying projective drawings' symbolic messages. According to Machover, "In some sense the figure drawn is the person and the paper corresponds to the environment" (1949, p. 35). Machover (1949) described the body in this theory as the most "intimate point of reference in any activity" (p. 5) and is a physical representation of the major inner tensions of the body image (Katz, 1951).

Few research studies reviewed by Swenson (1957) prior to 1957 supported the body image hypothesis. In fact, his review of five studies resulted in the conclusion that "definitive research on the basic meaning or significance of human figure drawings is lacking" (p. 437). Considered complex and difficult to measure, Swenson (1968) asked, "Is the body image a photograph, a verbal self-description, or a function of the interaction between a person's physical appearance and his self-concept?.....or a combination of something else?" (p. 23; cited in Knoff, 1986). Swenson (1968), Roback (1968), and Kahill's (1984) review of 31 body image related studies have yielded mixed results. Varied methodological approaches characterize recent studies of obese, mentally retarded, physically handicapped, and dermatological samples (Nathan, 1973; Wysocki & Wysocki, 1973; Wysocki & Whitney, 1965; Johnson & Greenberg, 1978). Kahill (1984) described the empirical evidence regarding the body image hypothesis as "mixed, with slightly more findings failing to support the hypothesis than supporting it" (p. 271).
Koppitz's Work

In Koppitz's (1968) book *Psychological Evaluation of Children's Human Figure Drawings*, she introduced methods for analyzing HFDs both as a measure of developmental maturity and as a projective technique. She hypothesized that structure, determined by age and maturation, and style of drawings reflect anxieties, attitudes, and concerns of the moment. According to Koppitz (1968), observing participants while they construct their HFDs is both revealing and essential to projective interpretation. It was hypothesized that those who drew the hands or feet of their figures before the head tended to experience difficulty fostering appropriate interpersonal relationships. A low self-concept, for example, was revealed if a participant could not complete a drawing representative of oneself.

According to Koppitz (1968, p. 75), "Youngsters do not necessarily identify with the person they draw, but instead, tend to depict the person who concerns them most at the time they make the HFD." While most children draw figures representing themselves, the quality and content of their drawings generally reflects underlying attitudes, feelings, anxieties, and conflicts regardless of whom the HFD represents. It is considered diagnostically useful to gather a general impression of the drawing without concern for detail due to the "need to grasp the tone and wholeness of the message conveyed in the drawing. The first impression helps to give a sense of the youngster's "feeling response.....emotional maturity and his or her balance or lack thereof" (Bolander, 1977, p. 64).
Koppitz (1968) analyzed human figure drawings for the presence or absence of 30 emotional indicators and believed that diagnostic significance is increased by considering their presence collectively as opposed to individually. Specifically, Koppitz (1984) defines emotional indicators according to three criteria: (a) they must differentiate between drawings of healthy and emotionally disturbed children; (b) frequency of occurrence in the nondisturbed population should be low (<16%); and (c) independent of maturation and age. The list of 30 emotional indicators were divided into three categories: (1) Quality Signs (Poor integration of parts, Shading of face, Shading of hands and/or neck, Gross asymmetry of limbs, Slanting figures, Tiny figure, Big figure, and Transparencies; (2) Special Features (Tiny head, Crossed eyes, Teeth, Short arms, Long arms, Arms clinging to body, Big hands, Hands cut off, Legs pressed together, Genitals, Monster or grotesque figure, Three or more figures spontaneously drawn, and Clouds; and (3) Omissions (No eyes, nose, mouth, body, arms, legs, feet, and neck) (Koppitz, 1968).

Koppitz and other authors (Hammer, 1958; Klopf & Taulbee, 1976; Mortensen, 1984) suggest that the true meaning or relevance of an emotional indicator can only be determined from the evaluation of the entire human figure drawing and personality assessment battery, taking into account the participant's age, cognitive ability, developmental/social background, and circumstances surrounding the production of the drawing.

Koppitz explored the clinical utility of 30 emotional indicators on human figure drawings of children five through 12 and found, following the comparison of HFDs of 76
clinic patients and 76 normal children, the occurrence and incidence of emotional indicators more often on clinic children's drawings than the well-adjusted children. A replication study by Fuller, Preuss, and Hawkins (1970) supported Koppitz's notion that emotional indicators occur more often in HFDs of disturbed children. Further review, however, suggests the number of emotional indicators cannot adequately differentiate between disturbed and normal groups because of agreement on only four items as to which ones were important. According to these authors, the emotional indicators are of little value in differentiating between disturbed and normal children when zero or one indicator are present but may helpful if there are more than two indicators present.

Koppitz (1966) also examined the emotional indicators depicted in drawings of shy and aggressive children. She found that shy children tended to draw cut off hands and tiny figures as opposed to aggressive children who drew long arms, teeth, big hands and gross asymmetry of limbs. While these studies demonstrated some evidence for the validity of cluster item interpretation of emotional indicators, a replication study by Lingren (1971) failed to identify significant differences between drawings of shy and aggressive children due to the limited description of subjects in the original study.

Like her predecessors, Koppitz's scoring system lacked a national standardization sample, experimental evidence of effectiveness, and poor documented reliability (McNeish & Naglieri, 1993). A number of researchers explored the clinical utility of the 30 emotional indicators in a variety of studies and found no empirical support for the Koppitz (1968) scoring system (Snyder & Gaston, 1970; Dieffenbach, 1977; & Mortensen, 1984).
Schildkrout, Shenker, and Sonnenblick's Work

Schildkrout, Shenker, and Sonnenblick's (1972) publication of Human Figure Drawings in Adolescence represents the first text to exclusively investigate psychopathology depicted in drawings of 12- to 19-year-olds. Selected from over 1,500 HFDs of sixth graders, college students, and in-patient adolescents, the drawings presented in the text demonstrate the range in the variety and maturity of adolescent human figure drawings.

Based on psychoanalytic theory and reliance on holistic and qualitative characteristics of drawings, the authors found HFDs to be helpful diagnosing psychopathology of all types but not personality disorders. Schildkrout, et al., found drawings useful when investigating adolescents' feelings about their bodies and sexuality, attitudes toward physical illness, and the reflection of neurological deficits, neurosis, and possible acting-out/suicidal behavior. However, no formal scoring system or empirical support for these hypotheses are offered.

Naglieri, McNeish, and Bardos' Work

Recently, a holistic approach to human figure drawing interpretation, Draw A Person: Screening Procedure for Emotional Disturbance (DAP:SPED) has been provided (Naglieri, McNeish, & Bardos, 1991). According to these authors, it is the accumulation of a number of signs and not individual signs on the human figure drawing that may provide a more accurate understanding of a child or adolescent's overall level of
emotional functioning. Their scoring system was designed to increase reliability, reduce the number of ambiguous items, and provide actuarial and normative scores.

According to Faust and Ziskin (1988), "Although more than 100 studies demonstrate the superiority of actuarial data over clinical judgment, few experts rely strictly on actuarial procedures; indeed, many do not even know that such methods exist" (p. 33).

Haddad (1995) considers the DAP:SPED a "paradigm shift away from the projective interpretation of HFDs, from high-level inference with regard to single signs in HFDs, from attributing each single sign in HFDs to some abnormal condition, and from the expert generating statements of abnormality with regard to individual signs in HFDs." This shift is based in part on findings by McNeish (1989) against continued projective interpretation of HFDs due to lack of empirical evidence.

The authors of the DAP:SPED (Naglieri, McNeish, & Bardos, 1991) have provided both a rationale and positive evidence for its use. Specifically, Naglieri and Pfeiffer (1992) compared the DAP:SPED scores of fifty-four normal students and fifty-four students identified as conduct and oppositional defiant disorders who attended a psychiatric day treatment facility. The DAP:SPED mean T-score earned by the 54 subjects in the clinical sample ($M = 56.63, SD = 10.27$) was significantly higher ($t = 4.05, p < .01$) than that of the 54 normal subjects ($M = 49.37, SD = 8.68$), indicating that more signs associated with emotional disturbance were exhibited by the clinical group than the normal group. A significant chi-square, $\chi^2 = 7.96 (1, N = 108), p < .005$ was
obtained indicating that DAP:SPED was effective in the identification of those requiring additional evaluation or intervention (Naglieri & Pfeiffer, 1992).

McNeish and Naglieri (1993) analyzed group differences for 81 regular education students and a matched sample of 81 students in special education placements for students with serious emotional disturbance. The mean T-score earned by the students with emotional disturbance ($M = 55.3$, $SD = 10.6$) was significantly higher ($t = 4.0$, $p < .001$) than that earned by the non-disturbed sample ($M = 49.5$, $SD = 8.6$). The d-ratio for the difference between means (divided by the pooled variance) was a medium .60 (Cohen, 1988, pp. 24-27). DAP: SPED scores, therefore, were moderately correlated with group membership ($r_{pb} = .36$, 95% confidence interval: .22 to .49 (McNeish & Naglieri, 1993).

**Item vs. Global Level Interpretation**

Human figure drawings continue to be used in the assessment, placement, and treatment decisions of children and adolescents despite considerable controversy regarding their validity. Relating specific indicators in drawings with specific diagnostic categories has been difficult for researchers where high false negative and high false positives rates have been reported (Dalby & Vale, 1977; Lingren, 1977; Mortensen, 1984).

Swensen (1957) reviewed empirical studies from 1949 to 1956 and concluded that research failed to support Machover's hypotheses regarding item level interpretation. In fact, he concluded that the global measures tended to yield higher inter-rater reliabilities.
Roback's (1968) review of empirical studies from 1949 to 1967 mirrors Swenson's (1957) findings failing to support Machover's hypotheses of item level interpretation. Kahill (1984) reports from her 1967 to 1982 literature review that global measures seem to be able to differentiate groups more successfully than structural and formal aspects of subjects' drawings.

**Human Figure Drawings and Sex Offenders**

A review of the literature on drawings of sex offenders revealed few empirical data and still fewer psychodiagnostic data. Further, the younger molester, considered a juvenile by the criminal justice system, is virtually ignored. Wysocki and Wysocki (1977) studied a group of male adult offenders ranging in age from 19 to 53 years and identified draw a person aspects assumed to have certain psychological implications in the area of sexual identification and disturbances, aggression, insecurity, narcissism, evasiveness, guilt, anxiety, and possible psychosis. According to Wysocki and Wysocki (1977), the following items suggest aggressive tendencies: slash mouth, stick fingers, talons, pointy foot, visible teeth, monster like-grotesque figure, clenched fist, and omission of arms. Items of sexuality are denoted by: long nose, long feet, button emphasis on fly area, waistline emphasis, omission of trunk, blank/hidden eyes, and hidden/missing hands.
Characteristics of Juvenile Delinquents with Sex Offense

Since Longo's (1982) discovery that adult male sex offenders started actually out sexually in adolescence, more clinical attention has been devoted to juvenile delinquent sex offenders. Perhaps this was due to the belief that few sex offenses of real consequence are actually committed by adolescents (Becker & Abel, 1984) and simply misunderstood as acts of sexual experimentation (Pierce & Pierce, 1987).

Contrary to this belief are accounts of rising crime statistics among youth, particularly sex offenses. During the latter 1970s, 30% of the rapists arrested were juveniles (Fehrenbach, Smith, Monastersky, & Deisher, 1986) and according to FBI statistics reported by Fagan and Wexler (1988), more than 50% of rapists were not yet twenty five years old. Since 1986, 19% of forcible rapes and 18% of all other sex crimes were committed by males under 19 (Blaske, Bordin, Henggeler, & Mann, 1989). In Ohio, for example, "commitments for sex offenses alone have increased a dramatic 60% since 1990" (Natalucci-Persichetti, 1996, p. 4).

Due to victims' reluctance to get involved with the criminal justice system and because the perpetrator usually is not apprehended with initial infractions of the law, Margolin (1979) suggests that estimates of sex crimes are probably underreported. Further, Groth and Loredo (1981) suggest that families are reluctant to report this type of crime because the offender is typically so young and known to the family and may consider the act experimentation or exploration without need for criminal prosecution. Amir (1971) estimated that fewer than half of all rapes are reported to police, and only 35% of children report sexual victimization to anyone (Finkelhor, 1979).
Numerous variables have emerged in the research of juvenile delinquent sex offenders' profiles including family discord, self-esteem and social issues, and knowledge about sexual relationships. Behavior and attitudes of parents were found to influence juvenile's sexually acting out behavior (McCord, McCord, & Verden, 1962) particularly when drug abuse, psychiatric disturbance, and marital discord are involved (Saunders & Awad, 1991; Kourany, Martin, & Armstrong, 1979).

Such families generally demonstrate low positive affect and have exhibited serious neglect or intra-familial violence witnessed by approximately 41% of juvenile sex offenders compared to 15% of a matched control group of juvenile delinquents (Van Ness, 1984; Blaske, Borduin, Henggeler, & Mann, 1989).

Low self-esteem and feelings of powerlessness and inadequacy appear to characterize juvenile delinquents with sex offense (Davis & Leitenberg, 1987; Lombardo & DiGiorgio-Miller, 1988). Several characteristics are reported that appear to differentiate these youths from normally developing peers.

Shorr (1966) assessed 80 male juvenile child molesters and found that they are "typically lonely and socially isolated from peers, prefer the company of younger children, are naive and lack suitable sex education, frequently experience disturbed family relations, and over half had engaged in previous delinquent behavior including other sexual offenses." Groth (1977) reports "that among 26 male juvenile offenders convicted of either rape or child assault, about three-fourths had committed a previous sexual offense."
Further, 86% of his sample had previous sexual experiences; thus, the sexual offense was not the first interpersonal sexual experience in the offender's life and did not simply represent naive curiosity or experimentation. Numerous authors consider juvenile sex offenders' negative self-image and lack of self-pride the result of numerous emotional traumas and victimization (Ryan, Lane, & Lafond, 1988; Kahn & Lafond, 1988). According to Fehrenbach et al. (1986), approximately 32% of juvenile sex offenders had no friends with 65% experiencing significant social isolation.

According to various researchers, juvenile delinquents with sex offense assume the role of power and control over victims while babysitting due to the availability of the victims, who are generally known to the offender and the heightened power or control the perpetrator seeks through offending (Lombard & DiGiorgio-Miller, 1988). Further, due to the pressure and tension of growing up often experienced by juveniles, Kourany, Martin, and Armstrong (1979) suggest that juvenile offenders commit such offenses while babysitting as an escape from loneliness or troubled homes. A majority of juvenile delinquents with sex offense (79%) were found to experience chronic learning problems as well, with over half (67%) requiring some type of special education programming (Saunders & Awad, 1991).

The relationship between sexual knowledge and sexual offending has been investigated by numerous authors. Juvenile delinquents with sex offense tend to lack appropriate sex education and understanding of bodily functions and normal sexual
expression (Shoor, Speed, & Bartelt, 1966; Scavo & Buchman, 1990) due to either being victims of sexual abuse or witnesses to atypical sexual attitudes and behaviors (Ryan, Lane, Davis, & Issac, 1987).

**Age, Sex, and Relationship of the Victim to the Offender**

Most of the research cited between 1977 and 1986 suggests that younger children tend to be victims of male juvenile sex offenders. According to Deisher et al. (1982), 46% of the victims were less than 10 years old, 29% between age 10 to 19, and 25% age 20 or above. Wasserman and Kappel (1985) found over 66% of the victims were not yet ten years old. In a study by Fehrenbach et al. (1986), 62% of victims were under age 12 and 44% were six-years-old or younger.

Over 80% of victims of incarcerated male juvenile sex offenders were female (Groth, 1977; Longo, 1982). Similar findings are reported by Fehrenbach et al. (1986), Van Ness (1984), and Wasserman and Kappel (1985) who report female victims 72%, 68%, and 77% of the time. Groth (1977) found that when the victim is a child, the proportion of boys is higher. Specifically, while 87% of peer-age and older victims of juvenile sex offenders were females, just 69% of child victims were female. A study by Van Ness (1984) found 91% of juvenile and adult rape victims were female with 63% of child victims being male (cited in Digiorgio-Miller, 1993).

According to findings by Groth (1977) and Van Ness (1984), most victims of juvenile sex offenses are known by the perpetrator. Groth's (1977) study of an incarcerated sample found 48% of victims were not known to juvenile offenders, 26%
were acquaintances, 15% were friends, and 6% were relatives. Also, when victims were children, they were less likely to be strangers than when victims were adults. In 55% of the cases of an incarcerated sample, the offender possessed some knowledge of the victim (Van Ness, 1984). When non-incarcerated samples were reviewed, the tendency is alarming. Findings by Wasserman and Kappel (1985) suggest "that 20% of victims were in the immediate family of the offender, 20% were extended family members, 51% were friends or acquaintances, and 9% were strangers." Also, Fehrenbach, et al., (1986), report that "33% of the offenses involved a relative, 50% an acquaintance, and only 17% a stranger."

Comparing Juvenile Delinquents with Sex Offense with Juvenile Delinquents without Sex Offense

Record reviews and interviews of 34 juvenile delinquents with sex offense from a sample of 242 violent juvenile delinquents without sex offense were the means by which Fagan and Wexler (1988) compared sex offending and non-sex offending juvenile delinquents. They concluded that juvenile delinquents with sex offense tended to have fewer problems with drugs and alcohol, less gang activity, and lower levels of self-reported delinquency than other juvenile delinquents. More often than not, juvenile delinquents with sex offense observed familial and spousal violence, were victim’s sexual molestation and abuse, and appeared to experience greater social and sexual isolation (fewer girlfriends) than other juvenile delinquents. Despite their awareness of the law, juvenile delinquents with sex offense tended to experience difficulty exhibiting
appropriate restraints or behavioral controls than other juvenile delinquents. Lewis, Shankok, and Pincus (1981) report that juvenile delinquents with sex offense tended to be younger and demonstrated less symptomatology, perhaps due to early detection and treatment, at the time of their first arrest than other juvenile delinquents. According to Awad, Saunders, and Levene (1981), 46% of the juvenile delinquents with sex offense versus 17% of other juvenile delinquents report difficulty fostering and maintaining appropriate peer relationships. The juvenile delinquents with sex offense tended to have IQ scores lower (below 80) than other non-offending juvenile delinquents and did not exhibit the extent of unruly behavior reported by the parents of juvenile delinquents without sex offense. Porter (1990) studied the psychosocial (interpersonal behaviors and coping skills), psychological (self-concept and affect), and developmental (compulsive, non-satisfying sexual patterns learned early in life), characteristics of 10 juvenile delinquents with sex offense, 10 juvenile delinquents without sex offense, and 10 juvenile non-offenders and found that juvenile delinquents with sex offense tend to be a "unique group of juveniles who differ widely from other subject groups in their internal norms, external expression, and mental dimensions" (p. 68). The juvenile sex offenders' point of view seems to lack the reality orientation of the other groups and is apparently deficient in environmental feedback due to an inflated view of oneself (cited in Digiorgio-Miller, 1993).

The juvenile delinquent with sex offense seemed to suffer from unresolved familial issues, depersonalize victims by ignoring victims' cues, prefer fantasy to reality, and have difficulty fostering appropriate peer relationships. The juvenile delinquent
without sex offense appeared to be able to express emotions and acknowledge various self-doubts, dangers, and fears (Porter, 1990).

"Ohio's Department of Youth Services operates a reception and assessment center where youth committed by Ohio juvenile courts undergo a variety of psychological, educational and medical tests and reports the following:

* Half have been diagnosed with mental disorders, 19% of whom are suffering from major mental illnesses such as depression and psychotic disorders.
* More than 8 of 10 come from households with incomes below $10,000.
* More than 6 of 10 live with single parents.
* 90% have substance abuse problems with alcohol, marijuana, crack or heroin.
* Nearly 1 of every 4 youths have children of their own.
* Nearly half are the offspring of at least one teenage parent.
* 52% are black; 43% are white.
* More than half of all youth are victims of domestic violence.
* On average, these youth are at least two grade levels behind in school"

(Natalucci-Persichetti, 1996, p. 4).

Characteristics of Repeat Adult Sex Offenders

A recent study (Prentky, Knight, & Lee, 1997) suggests that many convicted child molesters continue to commit crimes of a sexual nature after being released from prison and provides an indication of which child molesters are likely to offend again. Specifically, the researchers report the following characteristics that predicted a child molester's likelihood of recommitting a sexual offense, including: 1) "a strong sexual
preoccupation with children, both in fantasy and in the amount of time they spent around them, 2) prior sexual offenses, and 3) unusual sexual behaviors such as fetishism, exhibitionism, and transvestism" (Prentky, Knight, & Lee, 1997). Feelings of inadequacy, poor self-worth, and conflicting familial relationships are also common.

The same researchers also report the number of sexual charges, not necessarily convictions, against 115 child molesters in the sample was 14% at three years after release from prison, 30% at 10 years after release from prison; 46% at 20 years after release from prison; and 52% at 25 years after release from prison. "Ironically, the most dangerous sex offenders are essentially vagabonds who have no permanent address and hence are immune to public notification, while those who may be making a good-faith effort to resume normal, law-abiding lives are most adversely affected," when it comes to public or notification laws such as "Megan laws" named after a 7-year-old New Jersey girl who was sexually assaulted and murdered in 1989 (Prentky, Knight, & Lee, 1997).

**Conclusion**

The literature reviewed in this chapter sought to highlight the history of assessing social and emotional functioning in children and adolescents using human figure drawings as well as define variables characteristic of juvenile delinquent sex offenders. This literature continues to be characterized by inconsistent and controversial findings due in part to clinical opinion as opposed to empirical study. Therefore, the intent of this study is to narrow the gap that exists in psychological knowledge regarding the
identification of the social and emotional problems experienced by persons who are juvenile delinquents, particularly since few studies have utilized quantified human figure drawing systems.
CHAPTER 3
METHODOLOGY

Participants

The participants were 154 males (78 juvenile delinquents with sex offense, 76 juvenile delinquents without sex offense) ranging in age from 13 to 17 years in grades 9 through 12. The number of juvenile delinquents without sex offense who have sexually offended others and not apprehended or incarcerated for unrelated crimes is not known. All committed Felony one through Felony four offenses (see definition of terms in Chapter 1) and were receiving special education services (learning disabled, mentally retarded, severe behavior handicap) as defined by State of Ohio and Federal U. S. Guidelines while wards of Ohio’s Department of Youth Services. A sub-sample of 154 males from the DAP: SPED standardization sample served as the control group with a 100% match on all variables including age (by year), race, and sex (See Table 1). Geographic regions represented by the control group include 68% from the Midwest, 12% from the South, and 20% from the West. The total juvenile delinquent sample resided in the state of Ohio.

Instrument

The Draw A Person: Screening Procedure for Emotional Disturbance (DAP: SPED) (Naglieri, McNeish, & Bardos, 1991) is a measure that is designed to aid in the
<table>
<thead>
<tr>
<th>Age (Years)</th>
<th>with</th>
<th>without</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sex Offense</td>
<td>Sex Offense</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Males</td>
<td>78</td>
<td>76</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td>51</td>
<td>49</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>14</td>
<td>19</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>26</td>
<td>25</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>26</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>7</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>White</td>
<td>42</td>
<td>41</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>Black</td>
<td>36</td>
<td>35</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of all Juvenile Delinquents, Juvenile Delinquents with Sex Offense, and Juvenile Delinquents without Sex Offense.
<table>
<thead>
<tr>
<th>Geographic Region</th>
<th>Controls</th>
<th></th>
<th>Juvenile Delinquents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>105 68</td>
<td>154 100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>19 12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>30 20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Geographic Regions where Controls (Standardization Sample) and Juvenile Delinquents were drawn.

Identification of children and adolescents who may have emotional problems. The 55 item instrument was designed to be brief in both administration and scoring time and incorporate the recommendations that appear in the research literature regarding the caveats of past DAP scoring systems. Four new items were added to the DAP: SPED for this study by the author after consulting with Dr. Naglieri (personal communication with J. Naglieri, December, 1996). This instrument requires the participant to produce drawings of a man, woman, and self. Each drawing is scored based on the presence or absence (0 or 1) of items considered potential indicators of emotional disturbance. DAP: SPED raw scores are converted to a T-score (mean of 50 and a standard deviation of 10).

The DAP: SPED provides a scoring system for the drawings of a man, woman, and self that: (a) is comprised of items that are objectively and easily scored; (b) is recently normed on a nationally representative standardization sample; and (c) is
effective as a device to identify children and adolescents who may have emotional problems. The system intends to be an objective approach to determine the frequency with which unusual items that may be considered indicators of emotional conflicts occur in normal versus exceptional populations. Specific items are considered potential signs of emotional disturbance because of their selection following an exhaustive review of the literature of those signs used and researched as indicators of emotional disturbance, are found to occur infrequently (less than 16% in standardization) among normal individuals, and show appropriate psychometric properties. The methods resulted in DAP: SPED items that differ from those included in cognitive and achievement measures which vary considerably in difficulty. Nevertheless, the DAP: SPED total raw scores were essentially normally distributed.

The DAP: SPED standardization is comprised of 2,355 individuals and is representative of the U. S. population in terms of age, sex, geographic region, race, SES, and ethnicity. The reliability coefficients (internal, inter-rater, intra-rater) of the DAP: SPED are within a range considered good by Anastasi (1988). McNeish (1989) found correlations of .94 and .91 (p<.001) under intra and inter-rater reliability conditions in a study of 25 drawings scored two times by the same and different raters, respectively. Internal reliability coefficients were found to be .76, .77, and .71 for the ages 6-8, 9-12, and 13-17, which comprise the normative groups. The authors report that several validity studies comparing experimental group (Ns = 81, 49, 58, and 54) means were significantly (p<.001) higher than matched normal samples (Ns = 81, 218, 262, and 262) and are noted in the DAP: SPED manual.
Two new scales were also developed for use in this study in addition to the DAP:SPED Scale. One was the Sex Scale consisting of seven items that were hypothesized from the research literature as sexual in nature (See Appendix A for list of these items with comments). The second scale was the Aggression Scale consisting of seven items that were hypothesized from the research literature as aggressive in nature (See Appendix B for list of these items with comments).

Procedure

All participants were administered the Draw A Person: Screening Procedure for Emotional Disturbance (DAP: SPED) (Naglieri, McNeish, & Bardos, 1991) to evaluate their social/emotional status as part of their triennial reevaluation under Ohio’s Guidelines for the Identification of Handicapped Students. Each participant produced drawings of a man, woman, and self. The total raw score on the DAP: SPED was multiplied by 55/59 and rounded to the nearest whole number and converted to a DAP:SPED T-score using Appendix A in the test manual (mean of 50 and standard deviation of 10). The two new scale totals were determined by summing the number of emotional indicators present (0 to 7) on each of the scales and used in the data analysis.

Data Analysis

Three scores were obtained for each of the participants in the study. The first score was the Total DAP: SPED T-Score followed by scores for the Sex and Aggression Scales. A F-test for independent samples was conducted to compare the following groups: 1) controls and all juvenile delinquents, 2) controls and juvenile delinquents with...
sex offense, 3) controls and juvenile delinquents without sex offense, and 4) juvenile
delinquents with sex offense and juvenile delinquents without sex offense.

As a result, 12 F-tests for independent samples were conducted. To control the
alpha level across the 12 F-tests at the .05 level, the Bonferonni correction formula was
used. Sensitivity and specificity analysis of the DAP: SPED Total Scale, the Sex Scale,
and the Aggression Scale was conducted using three standard deviations and derived by
calculating and plotting cumulative percentages for each group comparison. D-ratio
statistics (difference between the means divided by the pooled standard deviations)
(Becker, 1991) were calculated to assess the efficiency of these differences between the
various groups and scales used in this study.

**Sensitivity and Specificity Analysis**

Sensitivity and specificity analysis is used to determine a test’s criterion related
validity. Sensitivity, or true positive rate, refers to the proportion of cases correctly
identified by the test as disordered. Specificity, or true negative rate, refers to the
proportion of non-cases correctly identified by the test as normal (Verhulst & Koot,
an intrinsic quality of a test, are not absolute values, and will vary with the samples on
which they were based and with the critical values chosen” (p. 49)

in this study, sensitivity refers to the proportion of juvenile delinquents correctly
identified or true positives. Specificity refers to the proportion of controls (normals)
correctly identified or true negatives.
According to Robins (1985), “the sensitivity of a test for the presence of a psychiatric disorder was higher in a patient sample than in a general population sample. The specificity of a test for the presence of a disorder will be higher in general population samples than in patient samples” (p. 49).
CHAPTER 4

RESULTS

This study explored four group comparisons: 1) Controls vs. all Juvenile Delinquents, 2) Controls vs. Juvenile Delinquents with Sex Offense, 3) Controls vs. Juvenile Delinquents without Sex Offense, and 4) Juvenile Delinquents with Sex Offense vs. Juvenile Delinquents without Sex Offense. Three scales were used to make four group comparisons. These scales were the Total DAP:SPED T-Score Scale, the Sex Scale, and the Aggression Scale. Twelve independent sample F-tests were used to make four group comparisons on the three scales. To maintain an alpha level for significance at .05, the Bonferonni correction was made resulting in an alpha level of .004 (.05/12) required for significance between any two groups compared in this study.

Sensitivity and specificity analysis was conducted for all four group comparisons using the three scales: 1) Total DAP:SPED T-Score Scale, 2) Sex Scale, and 3) Aggression Scale. A cut-off score of 55 (.05 standard deviations above the mean and recommended by the authors) was used with all 12 sensitivity and specificity analysis. A $d$-ratio (effect size statistic) was calculated for each of the four group comparisons using the Total DAP:SPED T-Score Scale.
Controls vs. Juvenile Delinquents

The mean Total DAP:SPED T-Score for all juvenile delinquents was significantly higher than the mean score for the control group, $F(1, 306) = 197.22, \ p < .0001$. The mean raw scores and mean Total DAP: SPED T-Scores with standard deviations are presented in Table 3.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile Delinquents</td>
<td>154</td>
<td>18.29</td>
<td>4.74</td>
<td>58.11</td>
<td>7.99</td>
</tr>
<tr>
<td>Controls</td>
<td>154</td>
<td>11.28</td>
<td>4.04</td>
<td>46.25</td>
<td>6.78</td>
</tr>
</tbody>
</table>

Table 3: Raw Score and T-Score Means with Standard Deviations on the Total DAP:SPED for Controls and all Juvenile Delinquents.

The performance of all juvenile delinquents on the Sex Scale (long nose, long feet, waistline emphasis, blank/hidden eyes, hidden/missing hands, crotch erasure, crotch shading, shown in Appendix A) did not differ significantly from that of the control group, $F(1, 306) = 4.33, \ p < .0384$ when using the Bonferroni corrected alpha level of .004. The means and standard deviations are presented in Table 4.

The mean score for all juvenile delinquents was significantly higher than the mean score for the control groups on the Aggression Scale (slash mouth, large figure,
aggressive symbols, talons, pointy foot/feet, teeth, clenched fist, see Appendix B), $F(1, 306) = 54.27, p < .0001$. The means and standard deviations in raw scores are presented in Table 4.

<table>
<thead>
<tr>
<th>Group</th>
<th>Sex Scale</th>
<th>Aggression Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw Score</td>
<td>Raw Score</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Juvenile Delinquents</td>
<td>154</td>
<td>3.49</td>
</tr>
<tr>
<td>Controls</td>
<td>154</td>
<td>2.95</td>
</tr>
</tbody>
</table>

Table 4: Mean and Standard Deviation Scores for Controls and all Juvenile Delinquents on the Sex and Aggression Scales.

Total DAP:SPED T-Score Scale. The results of sensitivity and specificity analysis with the Total DAP:SPED T-Scores for the controls and all juvenile delinquents using various cut-off scores are presented in Figure 1. When a cut-off score of 55 (.5 standard deviations above the mean) is used, the Total DAP:SPED T-Score correctly classified 55% of the juvenile delinquents and 93% of the controls. Using a cut-off score of 60 (1 standard deviation above the mean), 30% of the juvenile delinquents and 99.7% of the controls were correctly classified. With a score of 70 (2
standard deviations above the mean), 5% of the juvenile delinquents and 100% of the controls were correctly classified.

![Graph showing percentage against cut-off scores](image)

**Figure 1.** Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Controls and all Juvenile Delinquents.

**Sex Scale.** The results of sensitivity and specificity analysis using the Sex Scale for the controls and juvenile delinquents using various cut-off scores are presented in Figure 2. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 25% of the juvenile delinquents and 82% of the controls were correctly classified. When a cut-off score of 60 (1 standard deviation above the mean) is used, 15% of the juvenile delinquents and 92% of the controls were correctly classified. With a score of 70 (2 standard deviations above the mean), 5% of the juvenile delinquents and 98% of the controls were correctly classified.

42
Figure 2. Sensitivity and Specificity Analysis of the Sex Scale for Controls and all Juvenile Delinquents.

Aggression Scale. The results of sensitivity and specificity analysis using the Aggression Scale with controls and all juvenile delinquents using various cut-off scores and presented in Figure 3. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 34% of the juvenile delinquents and 90% of the controls were correctly classified. When a cut-off score of 60 (1 standard deviation above the mean) is used, 19% of the juvenile delinquents and 95% of the controls were correctly classified. With a score of 70 (2 standard deviations above the mean), 4% of the juvenile delinquents and 100% of the controls were correctly classified.

Effect Size. The effect size for the difference between the controls and all juvenile delinquents using the Total DAP:SPED T-Score Scale in standard deviation units was 1.59. This is considered large.
Figure 3. Sensitivity and Specificity Analysis of the Aggression Scale for Controls and all Juvenile Delinquents.

Controls vs. Juvenile Delinquents with Sex Offense

The mean Total DAP:SPED T-Score for juvenile delinquents with sex offense was significantly higher than the mean Total DAP:SPED T-Score for the control group $F(1, 230) = 142.48, p < .0001$. The mean raw scores and mean Total DAP:SPED T-Scores with standard deviations are presented in Table 5.
<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile Delinquents with Sex Offense</td>
<td>78</td>
<td>18.29</td>
<td>4.74</td>
<td>58.08</td>
<td>7.77</td>
</tr>
<tr>
<td>Controls</td>
<td>154</td>
<td>11.28</td>
<td>4.04</td>
<td>46.25</td>
<td>6.78</td>
</tr>
</tbody>
</table>

Table 5: Raw Score and T-Score Means with Standard Deviations on the Total DAP:SPED for Controls and Juvenile Delinquents with Sex Offense.

No significant difference was noted on the Sex Scale between the controls and juvenile delinquents with sex offense, $F(1, 230) = 2.17, p < .1425$ (See Table 6).

However, the performance of juvenile delinquents with sex offense on the Aggression Scale was significantly higher than for the control group, $F(1, 230) = 20.58, p < .0001$. The mean raw scores and standard deviations are presented in Table 6.
<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Sex Scale Raw Score</th>
<th></th>
<th>Aggression Scale Raw Score</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Juvenile Delinquents with Sex Offense</td>
<td>78</td>
<td>3.42</td>
<td>2.75</td>
<td>3.51</td>
<td>1.80</td>
</tr>
<tr>
<td>Controls</td>
<td>154</td>
<td>2.95</td>
<td>2.07</td>
<td>2.34</td>
<td>1.88</td>
</tr>
</tbody>
</table>

Table 6: Mean and Standard Deviation Scores for Controls and Juvenile Delinquents with Sex Offense on the Sex and Aggression Scales.

Total DAP:SPED T-Score Scale. The results of sensitivity and specificity analysis with the Total DAP:SPED T-Scores for the controls and the juvenile delinquents with sex offense using various cut-off scores are presented in Figure 4. When a cut-off score of 55 (.5 standard deviations above the mean) is used, the Total DAP:SPED T-Score correctly classified 50% of the juvenile delinquents with sex offense and 93% of the controls. Using a cut-off score of 60 (1 standard deviation above the mean), 23% of the juvenile delinquents with sex offense and 99.7% of the controls were correctly classified. With a score of 70 (2 standard deviations above the mean) 4% of the juvenile delinquents with sex offense and 100% of the controls were correctly classified.
Figure 4. Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Controls and Juvenile Delinquents with Sex Offense.

Sex Scale. The results of sensitivity and specificity analysis using the Sex Scale with controls and juvenile delinquents with sex offense using various cut-off scores are presented in Figure 5. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 27% of the juvenile delinquents with sex offense and 82% of the controls were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 17% of the juvenile delinquents with sex offense and 92% of the controls were correctly classified. With a score of 70 (2 standard deviations above the mean), 8% of the juvenile delinquents with sex offense and 98% of the controls were correctly classified.
Figure 5. Sensitivity and Specificity Analysis of the Sex Scale for Controls and Juvenile Delinquents with Sex Offense.

Aggression Scale. The results of sensitivity and specificity analysis using the Aggression Scale with the controls and juvenile delinquents with sex offense using various cut-off scores are presented in Figure 6. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 20% of the juvenile delinquents with sex offense and 82% of controls were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 10% of the juvenile delinquents with sex offense and 92% of the controls were correctly classified. With a score of 70 (2 standard deviations above the mean), 1% of the juvenile delinquents with sex offense and 98% of the controls were correctly classified.
**Effect Size.** The effect size for the difference between the controls and the juvenile delinquents with sex offense using the Total DAP:SPED $T$-Score Scale in standard deviation units was 1.66. This is considered large.

![Sensitivity and Specificity Analysis](image)

**Figure 6.** Sensitivity and Specificity Analysis of the Aggression Scale for Controls and Juvenile Delinquents with Sex Offense.

**Controls vs. Juvenile Delinquents without Sex Offense**

The mean Total DAP:SPED $T$-Score for juvenile delinquents without sex offense was significantly higher than the mean $T$-Score for the control group, $F (1, 228) = 61.14$, $p < .0001$. The mean raw scores and mean Total DAP:SPED $T$-Scores with standard deviations are presented in Table 7.
<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Raw Score</th>
<th>T-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Juvenile Delinquents without Sex Offense</td>
<td>76</td>
<td>18.31</td>
<td>4.88</td>
</tr>
<tr>
<td>Controls</td>
<td>154</td>
<td>11.28</td>
<td>4.04</td>
</tr>
</tbody>
</table>

Table 7: Raw Score and T-Score Means with Standard Deviations on the Total DAP:SPED for Controls and Juvenile Delinquents without Sex Offense.

The performance of the juvenile delinquents without sex offense on the Sex Scale did not differ significantly from that of the control group, $F (1, 228) = 4.25, p < .0404$.

The performance of the juvenile delinquents without sex offense on the Aggression Scale was significantly higher than the control group, $F (1, 228) = 61.14$, $p < .0001$. The mean raw scores with standard deviations are presented in Table 8.
<table>
<thead>
<tr>
<th>Group</th>
<th>Sex Scale</th>
<th>Aggression Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Juvenile Delinquents without Sex Offense</td>
<td>76</td>
<td>3.55</td>
</tr>
<tr>
<td>Controls</td>
<td>154</td>
<td>2.95</td>
</tr>
</tbody>
</table>

Table 8: Mean and Standard Deviation Scores for Controls and Juvenile Delinquents without Sex Offense on the Sex and Aggression Scales.

**Total DAP:SPED T-Score Scale.** The results of sensitivity and specificity analysis with the Total DAP:SPED T-Score for the controls and juvenile delinquents without sex offense using various cut-off scores are presented in Figure 7. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 45% of the juvenile delinquents without sex offense and 93% of the controls were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 34% of the juvenile delinquents without sex offense and 99.7% of the controls were correctly classified. With a score of 70 (2 standard deviations above the mean), 5% of the juvenile delinquents without sex offense and 100% of the controls were correctly classified.
Figure 7. Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Controls and Juvenile Delinquents without Sex Offense.

Sex Scale. The results of sensitivity and specificity analysis using the Sex Scale with the controls and juvenile delinquents without sex offense using various cut-off scores are presented in Figure 8. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 24% of the juvenile delinquents without sex offense and 82% of the controls were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 13% of the juvenile delinquents without sex offense and 92% of the controls were correctly classified. With a cut-off score 70 (2 standard deviations above the mean), 2% of the juvenile delinquents without sex offense and 98% of the controls were correctly classified.
Figure 8. Sensitivity and Specificity Analysis of the Sex Scale for Controls and Juvenile Delinquents without Sex Offense.

**Aggression Scale.** The results of sensitivity and specificity analysis using the Aggression Scale with controls and juvenile delinquents without sex offense using various cut-off scores are presented in Figure 9. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 47% of the juvenile delinquents without sex offense and 82% of the controls were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 29% of the juvenile delinquents without sex offense and 92% of the controls were correctly classified. With a cut-off score of 70 (2 standard deviations above the mean), 9% of the juvenile delinquents without sex offense and 98% of the controls were correctly classified.
Effect Size. The effect size for the difference between the controls and the juvenile delinquents without sex offense in standard deviation units was 1.63. This is considered large.

![Graph showing Sensitivity and Specificity Analysis of the Aggression Scale for Controls and Juvenile Delinquents without Sex Offense.](image)

**Figure 9.** Sensitivity and Specificity Analysis of the Aggression Scale for Controls and Juvenile Delinquents without Sex Offense.

**Juvenile Delinquents with Sex Offense vs. Juvenile Delinquents without Sex Offense**

No significant difference was noted between the mean Total DAP:SPED T-Score of juvenile delinquents with sex offense and juvenile delinquents without sex offense, $F (1, 152) = .00, p < .9582$. The mean raw scores and mean Total DAP:SPED T-Scores with standard deviations are presented in Table 9.
<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Raw Score</th>
<th>T-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Juvenile Delinquents</td>
<td>78</td>
<td>18.27</td>
<td>4.64</td>
</tr>
<tr>
<td>with Sex Offense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juvenile Delinquents</td>
<td>76</td>
<td>18.31</td>
<td>4.88</td>
</tr>
<tr>
<td>without Sex Offense</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Raw Score and T-Score Means with Standard Deviations on the Total DAP:SPED for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense.

No significant difference was noted on the Sex Scale between juvenile delinquents with sex offense and juvenile delinquents without sex offense, $F(1, 152) = .11, p < .7448$. However, the performance of juvenile delinquents without sex offense had significantly higher scores on the Aggression Scale, $F(1, 152) = 10.31, p < .0016$ than the juvenile delinquents with sex offense. The mean raw scores and standard deviations are presented in Table 10.
<table>
<thead>
<tr>
<th>Group</th>
<th>Sex Scale</th>
<th>Aggression Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Juvenile Delinquents with Sex Offense</td>
<td>78</td>
<td>3.42</td>
</tr>
<tr>
<td>Juvenile Delinquents without Sex Offense</td>
<td>76</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Table 10: Mean and Standard Deviation Scores for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense on the Sex and Aggression Scales.

Total DAP:SPED T-Score Scale. The results of sensitivity and specificity analysis with the Total DAP:SPED T-scores for the juvenile delinquents with sex offense and juvenile delinquents without sex offense using various cut-off scores are presented in Figure 10. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 51% of the juvenile delinquents with sex offense and 76% of the juvenile delinquents without sex offense were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 23% of the juvenile delinquents with sex offense and 87% of the juvenile delinquents without sex offense were correctly classified. With a cut-off score of 70 (2 standard deviations above the mean), 4% of the juvenile delinquents with sex offense and 98% of the juvenile delinquents without sex offense were correctly classified.

56
Figure 10. Sensitivity and Specificity Analysis of the Total DAP:SPED T-Score Scale for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense.

**Sex Scale.** The results of sensitivity and specificity analysis using the Sex Scale with juvenile delinquents with sex offense and juvenile delinquents without sex offense using various cut-off scores are presented in Figure 11. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 27% of the juvenile delinquents with sex offense and 75% of the juvenile delinquents without sex offense were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 17% of the juvenile delinquents with sex offense and 87% of the juvenile delinquents without sex offense were correctly classified. With a cut-off score of 70 (2 standard deviations above the mean), 8% of the juvenile delinquents with sex offense and 98% of the juvenile delinquents without sex offense were correctly classified.
Figure 11. Sensitivity and Specificity Analysis of the Sex Scale for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense.

Aggression Scale. The results of sensitivity and specificity analysis using the Aggression Scale with juvenile delinquents with sex offense and juvenile delinquents without sex offense using various cut-off scores are presented in Figure 12. When a cut-off score of 55 (.5 standard deviations above the mean) is used, 40% of the juvenile delinquents with sex offense and 53% of the juvenile delinquents without sex offense were correctly classified. Using a cut-off score of 60 (1 standard deviation above the mean), 10% of the juvenile delinquents with sex offense and 71% of the juvenile delinquents without sex offense were correctly classified. With a cut-off score of 70 (2 standard deviations above the mean), 1% of the juvenile delinquents with sex offense and 91% of the juvenile delinquents without sex offense were correctly classified.

Effect Size. The effect size for the difference between the juvenile delinquents with sex offense and juvenile delinquents without sex offense in standard deviation units was .007. This is considered small.
Figure 12. Sensitivity and Specificity Analysis of the Aggression Scale for Juvenile Delinquents with Sex Offense and Juvenile Delinquents without Sex Offense.
CHAPTER 5
DISCUSSION

This study compared the performance of juvenile delinquents with and without a history of sex offense to a control group using the DAP: SPED. Two additional scales (Sex and Aggression) were also derived from the original DAP item pool. The participants in the study included 78 juvenile delinquents with sex offense, 76 juvenile delinquents without sex offense, and 154 normal (controls) juveniles matched on gender, age, and race. Data for the control group were obtained from the DAP: SPED national standardization sample. Overall, the results supported the utility of the DAP: SPED for differentiating normal juveniles from juvenile delinquents with and without a history of sex offense.

Group Differences on the Total DAP: SPED

The entire juvenile delinquent population and those with and those without sex offense scored significantly higher on the Total DAP: SPED than the control group. The DAP: SPED mean T-score of 58.11 for the entire juvenile delinquent population was significantly higher than the mean of 46.25 for the control group and the effect size for the difference between the means was large. The DAP: SPED mean T-score of 58.08 for
the juvenile delinquents with sex offense was significantly higher than the mean of 46.25 for the control group and the effect size was large. The DAP: SPED mean T-score of 58.14 for the juvenile delinquents without sex offense was significantly higher than the mean of 46.25 for the control group and similarly, the effect size was large. However, the DAP: SPED mean T-score of 58.08 for the juvenile delinquents with sex offense did not differ significantly from the mean of 58.15 for the juvenile delinquents without sex offense and the effect size was small. This suggests that the juvenile delinquent sample, regardless of offense, exhibited more signs of social and emotional problems in their DAP protocols than similar age participants in the control group.

These findings mirror the results of four validity studies reported in the DAP: SPED manual where significant T-score and large effect size differences differentiating children and adolescents from special education and psychiatric/clinical samples and normal samples (Naglieri, McNeish, & Bardos, 1991) were found. These results also support the findings of others (Tharington & Stark, 1990; Naglieri & Pfeiffer, 1992; McNeish & Naglieri, 1993) who recognize the effectiveness of global interpretation of human figure drawings and lend further support for the validity of DAP: SPED differentiating individuals with emotional problems from normal samples. Like Naglieri and Pfeiffer (1992) who compared the DAP: SPED scores of normal students and students identified as conduct and oppositional defiant disordered who attended a psychiatric day treatment facility, the disturbed group earned higher scores than the control group. This indicated that more signs associated with emotional difficulties were
exhibited by the clinical group than the normal or control group. Similarly, McNeish and Naglieri (1993) analyzed group differences for regular education students and a matched sample of students in special education placements for students with serious emotional disturbance. They also found that the clinical sample was significantly higher than the control sample on DAP: SPED scores. Additionally, the effect size for the difference between the means, expressed in standard deviation units, was medium.

**Group Differences on the Sex Scale**

The Sex Scale used in this study was derived by selecting four DAP: SPED items (crotch erasure, crotch shading, vacant eyes, hidden hands) and the addition of three items considered sexual in nature (waistline emphasis, long nose, long foot/feet) as described in the literature (Wysocki & Wysocki, 1977). This new scale did not differentiate between any of the groups (non-significant F-tests were found). These findings are inconsistent with those reported by Wysocki and Wysocki (1977) who provided evidence that waistline emphasis, long nose, and long foot/feet are suggestive of sexual disturbance. The results of this study do not support their use or interpretation as items of sexuality sensitive to group differentiation or, the groups do not differ on sexual issues to begin with, even though their behavior is different.
Group Differences on the Aggression Scale

The total juvenile delinquent population, including those with and those without sex offense differed significantly from the control group on the Aggression Scale. The juvenile delinquents without sex offense, however, scored significantly higher on the Aggression Scale than the juvenile delinquents with sex offense. This finding is noteworthy due to the perceived aggressive nature of sexual offenses and the apparent lack of utility using these items to differentiate between these two groups. It is not known, however, the number of juvenile delinquents with sex offense who were convicted of gross sexual imposition, perhaps a less aggressive charge, than those convicted of seemingly more aggressive sex offenses such as aggravated rape, sexual battery, or felonious sexual penetration.

Koppitz (1966) and Griffith and Lemley (1967) have found items similar to the aggression items used in this study effective differentiating aggressive and non-aggressive groups. In contrast, Feyh and Holmes (1994) were not able to differentiate 40 conduct disordered from 40 non-conduct disordered children and adolescents by the presence or absence of aggressive emotional indicators. Thus, in this context, the present findings suggest some utility for differentiating juvenile delinquents with sex offense and juvenile delinquents without sex offense from a control group with the aggression items used in this study.

Sensitivity and Specificity Analysis

The accuracy of identification of juvenile delinquents with and without sex offense and controls was examined using an analysis of sensitivity and specificity.
The results from the sensitivity and specificity analysis of the four group comparisons on the three scales indicated that the Total DAP: SPED T-Score Scale (59 items) best classified groups with a specificity rating of 93% for controls and a sensitivity rating of 55% for all juvenile delinquents when using the authors’ recommended DAP: SPED cut-off score of 55 (.5 standard deviations above the mean). The Sex and Aggression Scales (seven items per scale) were less effective identifying group membership (sensitivity and specificity) but also contained fewer items and decreased reliability.

These sensitivity and specificity results are similar to those reported in the DAP:SPED manual (Naglieri, McNeish, & Bardos, 1991), where good group or accuracy of identification results were found. Lending further support for use of a cut-off score of 55 are findings by Naglieri and Pfeiffer (1992) who report a specificity rating of 77.78% and a sensitivity rating of 48.15% when comparing disruptive behavior disordered children and adolescents and controls. Also, when the same cut-off score of 55 was used to compare individuals with emotional problems and controls, a specificity rating of 68% and sensitivity rating of 49% was found (McNeish & Naglieri, 1993). The results of these studies offer some support for the validity of the DAP: SPED differentiating clinical from normal samples.

According to Naglieri and Gottling (1992), “the lower the score used to identify an individual, the lower the specificity (the proportion of non-cases identified, or true negatives), and as the scores increased, so did the specificity values. Inversely, the higher the cut-off score, the lower the sensitivity (the proportion of cases correctly identified, or true positives) rate” (p. 8). When addressing instruments which measure
the presence of psychiatric disorders, Verhulst and Koot (1992) state that “such tests usually detect severe cases more readily than mild ones” (p. 49) and, “as is always the case, sensitivity can only be increased at the expense of specificity and specificity can only be increased at the expense of sensitivity” (p. 52).

Conclusion

Though popular among clinicians for use evaluating intellectual and emotional development of children and adolescents, human figure drawing tests have been plagued by a history of poor objectivity, reliability, validity, and effectiveness differentiating those with emotional disorders from those without. The intent of this study was to utilize a recently developed human figure drawing system (DAP: SPED) with demonstrated psychometric foundations, to investigate whether or not juvenile delinquents with sex offense and juvenile delinquents without sex offense differ from a normal (control) sample of adolescents using a standardized scoring system.

The results of this study, demonstrating significant mean score differences, large effect sizes, and modest group identification rates (sensitivity and specificity), tend to support Naglieri, McNeish, and Bardos’ (1991) claim for the effectiveness of DAP: SPED as a screening measure for gross levels of maladjustment but not for differential diagnosis. For example, with the exception of the aggression items used in this study, no significant mean score or effect size differences emerged between juvenile delinquents with sex offense and juvenile delinquents without sex offense. While intuitively one might expect these groups to differ, Ageton (1983), after a three year study exploring personality
characteristics of juvenile delinquents concluded, "First, all of the findings point to the fact that sexual-assault offenders are basically delinquent youth whose sexually assaultive behavior is but one among many delinquent behaviors, all of which may be explained by the same set of variables" (p.129).

This scoring approach continues to demonstrate success differentiating clinical from control populations using a standardized scoring system, objective scoring rules, an overall or global rating system emphasizing number rather than presence of signs, good reliability and validity, and modest sensitivity and specificity ratings.

Limitations of Study

1. This is a descriptive study with all the limitations of such a study. For example, there was no random assignment of subjects to any of the three groups of participants used in the study, thereby affording every member of the population an equal probability of being assigned to any of the groups. Therefore, any differences observed between these groups may be do to differences that would not be present if three new non-random samples were drawn. Finally, the potential misclassification of some of the subjects in this study could be problematic. Car theft, for example, may serve as the offense for which the juvenile delinquent without sex offense is incarcerated, but does not exclude him from a prior sex offense or plea to a lesser charge.
2. The juvenile delinquents with sex offense and juvenile delinquents without sex offense were incarcerated and enrolled in special education classes when they were selected for use in this study, therefore limiting the generalizability of these results to similar samples.

Suggestions for Future Research

1. The DAP: SPED plus the four new items used in this study need to be used in other research studies involving other groups of both similar and different groups, including adults to further assess its validity for use identifying gross levels of emotional/behavioral maladjustment.

2. Future descriptive studies need to have the groups matched on more variables and data that might be obtained from behavior rating scales such as the Achenbach or Devereux Scales of Mental Disorders or the Devereux School Form to assess concurrent validity.

3. Studies need to be designed that are not restricted to a population that only includes participants enrolled in special education programs or are incarcerated, such as groups of non-aggressive youth or those with internalizing disorders such as depression or anxiety.

4. Experimental studies should be conducted exploring the value of this screening measure as a diagnostic tool in psychotherapeutic treatment programs that have been developed to reduce sexual offense recidivism.
REFERENCES


APPENDIX A
SEX SCALE
Items included in the Sex Scale

1. “Crotch Erasure is scored if erasure is apparent in the area of the figure’s crotch (below the waistline or belt and above the knee area of the leg)” (Naglieri, McNeish, & Bardos, 1991, p. 28).

2. “Crotch Shading is scored if pencil strokes are present on the figure’s crotch area (below the waistline or belt and above the knee area of the leg) which fill in an area by coloring or darkening (including stripes or checks on clothing)” (Naglieri, McNeish, & Bardos, 1991, p. 28).

3. “Vacant Eyes is scored if both the figure’s eyes (one if in profile) are empty (i.e., open circles)” (Naglieri, McNeish, & Bardos, 1991, p. 28).

4. “Hidden Hands is scored if the hands are hidden behind the back of the figure or in pockets” (Naglieri, McNeish, & Bardos, 1991, p. 29).

5. Waistline Emphasis is scored if at least two lines and/or shading in the waistline area is observed (Wysocki & Wysocki, 1977; Naglieri, McNeish, & Bardos, 1991, p. 6).

6. Long Nose is scored if this measurement item is >1 SD from the mean of the control group (Wysocki & Wysocki, 1977; Naglieri, McNeish, & Bardos, 1991, p. 6).

7. Long Foot/Feet is scored if this measurement item >1 SD from the mean of the control group (Wysocki & Wysocki, 1977; Naglieri, McNeish, & Bardos, 1991, p. 6).
APPENDIX B

AGGRESSION SCALE
Items included in the Aggression Scale

1. “Big Figure is scored if the figure exceeds both the vertical and horizontal dimensions of Box 3 (use template). Articles of clothing such as hats or shoes are included in the measurement, although other objects (e.g., handbag, briefcase, backpack, baseball bat) are not included” (Naglieri, McNeish, & Bardos, 1991, p. 27).

2. “Slash Mouth is scored if the figure’s mouth is a straight line or slash” (Naglieri, McNeish, & Bardos, 1991, p. 28).

3. “Teeth is scored if teeth are present in the figure’s mouth” (Naglieri, McNeish, & Bardos, 1991, p. 28.)

4. “Fists is scored if the hands are made into fists” (Naglieri, McNeish, & Bardos, 1991, p. 47).

5. “Talons is scored if one or more fingers are clearly pointed (like a claw) or knife-like” (Naglieri, McNeish, & Bardos, 1991, p. 29).

6. “Aggressive Symbols is scored for the presence of one or more aggressive symbols, gestures, or written statements (e.g., guns, knives, clubs, written profanity, or other symbols of aggression)” (Naglieri, McNeish, & Bardos, 1991, p. 29).

7. Pointy Foot/Feet is scored if one or both feet are clearly pointed (Wysocki & Wysocki, 1977).
APPENDIX C
EXAMPLES OF HUMAN FIGURE DRAWINGS
Figure 13: illustrates emotional indicators present in a human figure drawing of a juvenile delinquent with sex offense (crotch shading, vacant eyes, hidden hands, and waistline emphasis from the Sex Scale); other emotional indicators present include teeth and pointy feet (from the Aggression Scale); arms pressed to torso, and fingers omitted.
Figure 13
Figure 14: illustrates emotional indicators present in a human figure drawing of a juvenile delinquent without sex offense (big figure, talons, slash mouth from the Aggression Scale); other emotional indicators present include crotch shading and waistline emphasis (from the Sex Scale), tall figure, legs together, and long feet.
Figure 15: illustrates the human figure drawing of an adolescent from the control group (fists, feet shading, and pointy feet).