To the Graduate Council:

I am submitting herewith a dissertation written by Holly Michelle Hutchins entitled "An Evaluation of the Psychometric Properties of the Universal Nonverbal Inventory of Personality." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

R. Steve McCallum, Major Professor

We have read this dissertation
and recommend its acceptance:

Acceptance for the Council:

Vice Chancellor and Dean of Graduate Studies

Thesis
2005b
.H89

AN EVALUATION OF THE PSYCHOMETRIC PROPERTIES OF THE UNIVERSAL NONVERBAL INVENTORY OF PERSONALITY

A Dissertation Presented for the Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Holly Michelle Hutchins

August 2005

DEDICATION

This dissertation is dedicated to my parents, Mike and Mary Hutchins, my brother, Daniel Houston, and my grandmother, Dorothy Reed, for always believing in me.

Thank you, Mama, for always being my role-model and best friend. I cherish each of those late nights that we watched Animal Planet and MTV and ate junk food like teenagers. I pray that I grow up to be a mom like you. Thank you, Daddy, for always letting me stay your "little girl" and bringing me birthday cakes even when it wasn't my birthday. I pray that you know how much I love you. Thank you, Daniel, for always loving me and being proud of me. Words cannot express how much that has motivated me during these last few years. Thank you, Granny, for all the smiles, hugs, talks, and quilts. Your support has given me confidence to continue with my dreams.

I would also like to thank my boyfriend and best friend, Craig Morren, for asking me to be his girlfriend. I am grateful to have been able to share this journey with you.

I love each of you with all my heart.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my wonderful committee members, Dr. R. Steve McCallum, Dr. Sherry Bain, Dr. John Lounsbury, and Dr. Charles Hargis, for generously providing encouragement, support, and guidance. Each of them, in their own way, helped make this possible for me. A very special thank you goes to Dr. R. Steve McCallum, my committee chair, for making this a wonderful experience that I will never forget. Dr. McCallum is not only an excellent professor but also a great person. His enthusiasm for this project, patience, and support are greatly appreciated.

To Sara McCane and Stacy Bliss, thank you for helping with the data entry. Your help allowed me to concentrate on the writing of my dissertation. I am very proud of our work.

ABSTRACT

The Universal Nonverbal Inventory of Personality (UNIP; McCallum, in preparation), which is designed to measure psychopathological maladjustment of children and adolescents who have limited or no English language proficiency, and the Behavior Assessment System for Children-Second Edition Self-Report (BASC-II; Reynolds & Kamphaus, 2004) were administered to 100 participants to establish reliability and concurrent validity. Some of the UNIP scales yielded adequate reliability: Anxiety (.85), Depression (.84), Atypicality (.78), Academic Problems (.77), Attention Deficit Hyperactivity Disorder (.75), and Conduct Problems (.73); others possess only marginal reliability: Social Maladjustment (.64) and Consistency (.54) scales. Correlation coefficients between the UNIP and the BASC-II support the validity of the UNIP Depression, Anxiety, Attention Deficit Hyperactivity Disorder, Academic Problems, and Conduct Problems scales. Results provide only marginal support for the concurrent validity of the UNIP Social Maladjustment and Atypicality scales. The utility of the UNIP for clinicians and the need for future investigations are discussed.

TABLE OF CONTENTS

1. INTRODUCTION 1

Purpose 1

Purpose of Personality Assessment 1

Personality Assessment in Schools 5

Existing Nonverbal Personality Inventories 13

2. STATEMENT OF THE PROBLEM 17

3. RESEARCH QUESTIONS 19

4. METHODS 20

Participants 20

Instruments 20

Procedures 23

Data Analysis 24

5. RESULTS 26

Reliability of the UNIP 27

Validity of the UNIP 28

6. DISCUSSION 30

Internal Consistency 30

Concurrent Validity of the UNIP with the BASC-II 35

Summary and Implications 40

Limitations and Suggestions for Further Research 41

LIST OF REFERENCES 44

APPENDICES 52

Appendix A – Universal Nonverbal Inventory of Personality 53

Appendix B – UNIP Item Interpretation and Scale Assignment... 69

Appendix C – Administration of the UNIP 77

Appendix D – Tables..... 85

VITA..... 94

LIST OF TABLES

Table D1. Descriptive Statistics of the Universal Nonverbal Inventory of Personality 86

Table D2. Descriptive Statistics of the Behavior Assessment System for Children-Second Edition 87

Table D3. Internal Consistency Reliability Estimates (Cronbach's Alphas) of the Universal Nonverbal Inventory of Personality Scales 88

Table D4. Percentage of Inconsistent Responding to Universal Nonverbal Inventory of Personality Consistency Pairs 89

Table D5. Item-Total Correlations for the Universal Nonverbal Inventory of Personality Scales 90

Table D6. Correlation Coefficients between Universal Nonverbal Inventory of Personality Scales and Behavior Assessment System for Children-Second Edition Scales 93

1. INTRODUCTION

Purpose

Although many tests of personality exist, not one has been standardized using a nonverbal administration and response format designed to identify maladjustment (e.g., depression and anxiety) among children and adolescents who cannot use language to communicate. The Universal Nonverbal Inventory of Personality (UNIP; McCallum, in preparation) is a newly developed nonverbal personality measure designed to overcome the major limitations associated with conventional verbal-laden personality inventories. Despite its clinical appeal and face validity, there are no data yet available to support its psychometric integrity. Therefore, the purpose of this study is to investigate UNIP psychometric properties, including its reliability and validity. Specifically, this study was designed to evaluate internal consistency reliability as well as concurrent validity of the UNIP with the Behavioral Assessment System for Children, Second Edition-Self-Report of Personality for Adolescents (BASC-II SRP-A; Reynolds & Kamphaus, 2004).

Purpose of Personality Assessment

According to Aiken (1997), personality assessment refers to the measurement and evaluation of emotions, thoughts, attitudes, and behavioral traits that are reflective of one's personality by means of instruments and procedures (e.g., rating scales and observations). The goals of personality assessment and the methods chosen to reach those goals vary widely (Aiken, 1995). For instance, some personality assessment instruments are used to facilitate the identification and diagnosis of maladjustment and psychiatric disorders by distinguishing between psychologically "normal" and abnormal people in

clinical settings. Among the personality assessment instruments commonly administered in clinical settings are the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrm, Graham, Tellegen, & Kaemmer, 1989), the Rorschach Inkblot Test (Rorschach, 1921), the Thematic Apperception Tests (TAT; Morgan & Murray, 1935), the Beck Depression Inventory (BDI; Beck and Steer, 1987), the Bender-Gestalt Test (BGT; Bender, 1938), the Symptom Checklist-90R (SCL-90-R; Derogatis, 1994), the Million Clinical Multiaxial Inventory-II (MMCI-II; Million, 1987), and various other projective drawing and sentence completion tests (Aiken, 1997; Watkins, Campbell, Neiberding, & Hallmark, 1995). These types of personality instruments are typically used by psychiatrists and psychologists with patients who have some fairly serious problems (e.g., anxiety, depression, anger, impulsivity, stress, and suicidal thoughts) with which they are unable to cope (Aiken, 1997). In contrast, other personality tests are used to measure personality characteristics and adjustment in relatively "normal" individuals by uncovering their characteristic ways of feeling, thinking, and behaving through measure of attitudes, values, interests, and other psychosocial characteristics (Segal & Coolidge, 2003). This type of information is often used in military and government, educational, industrial/organizational, and health settings for purposes of selection, placement, classification, dismissal, promotion, and the rotation of trainees, employee, and students. Among the personality tests commonly used for these purposes are the Myers-Briggs Type Indicator (MBTI; Briggs & Myers, 1985), the Sixteen Personality Factor Questionnaire (16PF; Cattell, Eber, & Tatsuoka, 1970), the Edwards Personal Preference

Schedule (EPPS; Edwards, 1959), and the California Psychological Inventory (CPI; Gough, 1987) (Aiken, 1997).

Tests of personality can be further subdivided into two relatively exclusive categories: objective and projective. Objective personality tests typically are psychometrically sound and require standardized administration using clear, specific questions or items, for which respondents choose from a limited range of responses (e.g., answering "yes" or "no"). In comparison, projective personality tests typically require the administration of ambiguous stimuli and allow open-ended response formats (e.g., examinees are asked to tell a story based on a neutral picture) (Segal & Coolidge, 2003). Mental health practitioners typically use both projective and objective approaches. For example, it is commonplace for practitioners to use projective drawings, while also observing and unsystematically interpreting examinees' nonverbal behaviors (i.e., facial expressions and motor behaviors) during evaluations (Wasserman, 2003). However, treatment and diagnostic decisions are typically based on information collected from objective instruments since projective techniques are often criticized for a lack of commonly accepted standards of reliability, validity, and clinical utility (Lachar & LaCombe, 1983). Since projective techniques are time consuming to administer, score, and interpret, mental health practitioners often opt to use objective personality assessment procedures when there are no specific referral questions requiring the use of projective techniques. In addition projective techniques require considerable verbal interaction and cannot be used effectively with examinees who are language impaired or by those who

are not proficient in English. Similarly, objective tests typically require verbal directions, item content, and/or response format.

If the goal is to use either projective or objective measures to assess the personality structure or psychopathology for examinees with an absence, disorder, or delay in verbal communication, the choices are extremely limited and, of the nonverbal tests available, none have been standardized for use in the United States. Because the U.S. population is becoming increasingly diverse and because many native examinees are non-English speaking or have limited English proficiency, a psychometrically sound nonverbal instrument is needed. In the next section I review briefly the history of personality assessment, discuss current personality assessment practices in schools, describe two nonverbal objective personality tests that are available, and conclude with a description of the UNIP.

The earliest attempt to systematically and objectively measure and apply information gathered about an individuals' personality occurred during World War I when the first psychometric personality inventory, the Woodworth Personal Data Sheet (Woodworth, 1917), was developed for the purpose of military selection. This paper-and-pencil inventory consisted of statements which focused on psychoneurotic symptoms and was developed to serve as a psychiatric screener for soldiers. Although it was developed too late to be put into use during the war, it paved the way for the development of many personality measurements that are used today (Aiken, 1997). Specifically, the Woodworth Personal Data Sheet was devised to shed light on individuals' current state of psychological functioning. Subsequently, instruments were developed with this same

goal, and consequently, businesses and industries, government agencies, clinical and counseling facilities, and educational settings adopted and still use personality instruments such as the MBTI (Briggs & Myers, 1985) and the MMPI-2 (Butcher et al., 1989).

Personality Assessment in Schools

According to Wasserman (2003), psychometrically sound personality instruments can be used to evaluate individuals' current symptomatology, identify problems in functioning and living, uncover potential environmental supports and intrapsychic coping resources, and describe underlying interpersonal and characterological styles. While such information is valuable to individuals of all ages, information of this nature is of particular interest for mental health practitioners (i.e., school, counseling, or clinical psychologists, social workers, psychiatrists, guidance counselors) who work with children and adolescents. In particular, psychological assessments within school settings facilitate the collection of pertinent information that may explain possible causes for or contingencies related to students' behavioral, social-emotional, or affective problems (Knoff, 1983). Such information is helpful in understanding the significant number of social-emotional problems that students manifest in schools. By gaining a deeper understanding of what the child is experiencing internally, mental health practitioners can provide parents and teachers with recommendations that may decrease, resolve, or even prevent such problems as drug abuse, pregnancy, and suicide and can offer insight regarding the emotional impact of academic failure, rejection, poverty, and divorce (Knoff, 1983).

Within school settings, personality assessments are used for a variety of purposes, including screening, diagnosis, placement, and treatment. For example, schools often administer a personality test at the beginning of the school year for the purpose of identifying students who are or can be expected to experience psychological problems during the school year (Aiken, 1999). Often teachers refer individuals with significant emotional or behavioral problems that seem to be causing an adverse impact on their educational performance in the learning environment. Typically these referrals are made "when a child or adolescent's behavior problems, interactions, or ecological situation become so significantly disruptive that mental health intervention appears warranted" and/or special education services are deemed necessary for the student to progress within the educational setting (Knoff, 1986, p. 3). Information obtained from personality evaluations can be used to develop, implement, and evaluate interventions. According to Knoff (1995):

> While the ultimate personality assessment goal is to develop and implement effective intervention programs for referred students, other goals for the school psychologists might be (a) to determine who "owns" a specific referred problem (e.g., the referred child, a referring teacher or parent, a dysfunctional system, or a combination thereof); (b) to validate hypotheses explaining how a referred child's behaviors are being caused, encouraged, reinforced, or supported; (c) to create a sound baseline of data so that interventions can be evaluated from an appropriate context so that an accurate presenting history can be documented; and (d) to

identify the referred child's behavioral assets and the home and school's resources so that they might be integrated into an intervention program. (p.1281)

School district personnel primarily use personality assessment to help "determine a student's eligibility for special education services" (Knoff, 1983, p. 1281). Psychological assessments of children enrolled in schools became essential in 1975 when the Education of All Handicapped Children Act (PL 94-142), now known as the Individuals with Disabilities Education Act (IDEA), mandated the provision of appropriate educational opportunities for students with severe social-emotional problems (Prout, 1983). According to Lachar and LaCombe (1983), school psychologists use personality tests to satisfy a diagnostic function and "are routinely called upon to assess the nature of a child's psychopathology so that appropriate special services can be provided" (p. 399). As a result of IDEA, personality assessment became a major function of school psychologists (Prout, 1983).

School psychologists use personality measures in nearly half of their referral cases (Goh, Teslow, & Fuller, 1981). Procedurally, following an evaluation referral, school psychologists administer personality measures to obtain information that will become an integral part of a comprehensive individualized education plan used by a multidisciplinary child study team (Knoff, 1983). In many cases the information will be used to determine whether students with psychological problems meet specified criteria for a diagnosis of serious emotional disturbance (SED) and qualify for special education services under IDEA. According to IDEA, SED refers to

> a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child's educational performance: (A) An inability to learn that cannot be explained by intellectual, sensory, or health factors. (B) An inability to build or maintain satisfactory interpersonal relationships. (C) Inappropriate types of behavior or feelings under normal circumstances. (D) A general pervasive mood of unhappiness or depression. (E) A tendency to develop physical symptoms or fears associated with personal or school problems. [Code of Federal Regulation, Title 34, Section 300.7 (c)(4)(i)]

According to IDEA, SED also includes schizophrenia, but excludes children with social maladjustment, unless it is determined that they also have an emotional disturbance. Although states define SED differently and specify criteria to be used by local schools in the identification of emotionally disturbed children, school psychologists, regardless of the state in which they are employed, generally use personality instruments to help determine the presence and severity of emotional disturbances (i.e., see the Tennessee Department of Education's Special Education Manual (2003), which specifies that an individual evaluation of psychological strengths and weaknesses should include both personality and intelligence factors).

As illustrated above, personality measures are essential to the process of determining special education eligibility, especially when attempting to diagnose an emotional disturbance. Unfortunately, the current identification processes often result in youth going without special services. Literature indicates the SED category is the most

underrepresented of all the special education categories (Kidder-Ashley, Deni, Azar, & Aderton, 2000; Forness & Knitzer, 1992). It has been reported that the national identification rate for students with SED has remained stable at about 0.9% since 1976 when national data were first collected (Kidder-Ashley et al., 2000; Oswald & Coutinho, 1995). This means that less than 1% of the student population is being served under the category of SED. However, Katsiyannis, Landru, Bullock, and Vinton (1997) noted (as cited in Kidder-Ashley et al., 2000), "professional estimates suggest that this figure should be at least double, and may be as high as 6 to 10%" (p. 560). Even a conservative identification rate of 3-6% is suggested by Kaufman (1994) as being a more accurate estimate of emotionally disturbed students within schools. Although this range of estimates is large, there is consensus among researchers that many children with emotional problems are going unidentified, suggesting that the assessment procedures within school settings may be insufficient and additional screening/diagnostic techniques are needed.

To address the under-identification of emotionally disturbed students, the Center for Effective Collaboration and Practice (2001) reported research suggesting the current identification process is subjective, reactive, motivated by institutional needs, constrained by a lack of linguistically and culturally appropriate assessment tools, limited by parental concerns regarding pejorative labels (U.S. House of Representatives, 1997) and inappropriate placement, and a lack of collaboration among professionals and families (McInerney, Kane, & Pelivan, 1992). Kiddler-Ashley et al. (2000) suggest that the under-identification of emotionally disturbed students can often be attributed to the vague

terminology of the federal definition of SED, as well as the variations in definitions and criteria across states. While each of these shortcomings of the current identification process need to be addressed and resolved, the focus of this study is to address the customary reliance of mental health practitioners on verbally-laden personality tests and the lack of alternative culturally and linguistically appropriate personality measures for individuals who cannot use language to communicate or have limited English language proficiency and would benefit from nonverbal measures of personality. Specifically, according to Wasserman (2003), nonverbal personality assessment measures may be needed with

> (a) individuals with neurologically based acquired language disorders (e.g., aphasia, language-based learning disabilities), (b) individuals with varied cultural, linguistic, or national backgrounds (e.g., non-English speakers), (c) individuals who are illiterate or poorly educated, (d) individuals who are deaf or hard of hearing, (e) individuals with forms of emotional disturbance that are manifested through an inability or unwillingness to produce an adequate and unconstrained sample of verbal behavior (such as may be found in cases of severe depression, some psychoses, or selective mutism), or (f) individuals who are prone to misrepresent themselves on verbal self-report measures. (p. 284)

For these individuals, language no longer serves as a window into their psychological well-being, but instead as a barrier, and an alternative personality instrument would be both optimal and ethical.

Traditional verbally-laden personality measurements are typically designed for use with youth at least 13 years of age; test content is written at the sixth-grade reading level. Unfortunately, even a sixth-grade reading level may be inappropriate for many adolescents in need of psychological evaluations due to language-based learning difficulties or impairments that interfere with their reading abilities (Wasserman, 2003). In addition, such language-based inventories hinder evaluations of young children and examinees who have attention-span and/or psychiatric problems, and who may benefit from novel appealing pictorial stimuli (Paunonen, Jackson, & Keinonen, 1990).

Because the world population is rapidly changing there is an influx of immigrants into the United States; schools are becoming multicultural, multiracial, multilingual, and multiethnic (McCallum, Bracken, & Wasserman, 2001). Kindler (2002) (as cited in Ochoa, Riccio, Jimenez, Alba, & Sines, 2004) reported that 9.3% of the school-aged population is comprised of limited English proficient (LEP) students. Furthermore, the student population of English language learners (ELL) is increasing at a much faster rate compared to the general student population (Ochoa et al., 2004). According to the National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs (NCELA, 2002), during the 1990s, overall student enrollment in schools increased by 24.2%, whereas the ELL student enrollment dramatically increased by 105%. As a result, as cited in McCallum (2003), there are approximately 200 languages spoken by students who attend the Chicago City schools (Pasko, 1994), more than 140 languages are spoken across the state of California (Puente, 1998; Unz, 1997), more than 80 in Palm Beach County Schools (Fast Fact, 1996), more than 60 in the

schools of Plano, Texas (Power, 1996), and 61 in Knox County, Tennessee (Forrester, 2000). Many of these children do not read at the sixth-grade reading level or even understand spoken English. Even worse is that children who are ELLs are frequently faced with a variety of stressful issues associated with language differences and immigration that can result in a significant risk for mental health problems, which suggests that they are in a greater need for emotional disturbance evaluations (Ochoa et al., 2004).

Given the lack of bilingual mental health practitioners (i.e., school psychologists) and lack of personality instruments in native languages, the ethical alternative (to testing non-English speaking children in their native languages) is to omit language as a variable and employ the use of nonverbal measurements (Frisby, 1999). As a result, according to Ochoa and colleagues (2004), the most commonly used measures when conducting evaluations for ELLs are primarily nonverbal, yet projective in nature. The most commonly used instruments include the Bender Visual Motor Gestalt, Draw-A-Person, House-Tree-Person, and Kinetic Family Drawing, which require the examinee to understand a limited amount of instruction in English and require little to no verbal responses from the examinee. Unfortunately, these instruments are also among the least psychometrically sound personality tests. Therefore, there is a need for psychometrically sound objective nonverbal personality tests to use with examinees who cannot use language to communicate or have limited English language proficiency.

Existing Nonverbal Personality Inventories

The term "nonverbal assessment" is frequently used but often implies a variety of assessment procedures. For Bracken and McCallum (1998), who developed the Universal Nonverbal Intelligence Test (UNIT), nonverbal assessment characterizes an evaluation procedure in which there are no expressive or receptive language demands on either the examinee or the examiner. However, Wasserman (2003) holds that a test is nonverbal if it "involves a relatively brief verbal instructional set (and therefore makes limited demands on the examinee's receptive language) and requires little or no verbal response (thereby involving minimal expressive language) on the part of the examinee" (p. 284). For the purposes of this research, nonverbal assessment is operationally defined as a test that requires examiners to use a brief instructional set with gestures, demos, samples to convey task demands with limited demands on the examinee's receptive language, and requires no verbal responses of the examinee. Based on these criteria, a personality inventory is categorically nonverbal if it provides a simple set of directions, presents items that tap into an examinee's emotions, thoughts, attitudes and behavior traits via nonlinguistic pictorial stimuli, and requires no verbal responses of the examinee.

To date, there are very few objective and structured nonverbal self-rating and self-report tests of personality. In fact, there are only two viable nonverbal measures available, including the Nonverbal Personality Questionnaire (NPQ; Paunonen & Jackson, 1998) and Five-Factor Nonverbal Personality Questionnaire (FF-NPQ) (Paunonen, Ashton, & Jackson, 2000). The NPQ consists of 136 pictorial items intended to portray 16 of the traits depicted in Murray's (1938) system of needs. Such personality

traits, or needs, include Achievement, Affliction, Aggression, Autonomy, Dominance, Endurance, Exhibition, Impulsivity, Nurturance, Order, Play, Sentience, Social Recognition, Succorance, Thrill-seeking, and Understanding, which correspond directly to the personality traits that are measured by Jackson's (1984) Personality Research Form (PRF), which is a more established verbal personality measure (Paunonen et al., 1990; Paunonen & Ashton, 2002). The FF-NPQ consists of 60 items intended to assess the Big Five personality factors, which include Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Paunonen & Ashton, 2002). Most of the items on the FF-NPQ were selected from the NPQ. For both measures, the items are represented by stick-figure drawings of a central character performing a behavior in a specific situation.

In the self-report formats of the NPQ and FF-NPQ, respondents are presented a picture booklet and asked to look at each item and "rate the likelihood that [he or she] would engage in the type of behavior shown" (Paunonen & Jackson, 1998). These instructions are written in the booklet and require translation, as needed. All item responses are recorded by means of a 7-point Likert scale with one end of the scale labeled "extremely unlikely" and the other end labeled "extremely likely" with similar verbal anchors placed appropriately along the scale. The NPQ requires approximately 25-30 minutes to complete. The format of the FF-NPQ is similar, but due to its shorter length, respondents typically complete it in approximately 10 minutes. Scoring of the instruments is accomplished by hand.

Both the NPQ and FF-NPV demonstrates satisfactory psychometric properties. The internal consistency reliabilities of the 18 NPQ scales appear to be adequate, with mean coefficient alpha across scales of 0.75 for a Canadian sample, an average of 0.67 for four European samples, and 0.61 for a Chinese sample (Paunonen, Keinonen, Trzebinski, Forsterling, Grishenko-Roze, Kouznetsova, & Chan, 1996). The mean internal consistency reliability of the FF-NPQ scales appears to be satisfactory, with an average coefficient alpha of 0.80 for a Canadian sample (Paunonen, Ashton, & Jackson, 2001). The internal consistency reliabilities of the NPQ scales and the FF-NPQ scales compare favorably with those of the corresponding verbal PRF scales and NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992) scales respectively.

The convergent validity of the NPQ and FF-NPQ appears to be acceptable. Although there is variability across cultures, Paunonen, Keinonen, Trzebinski, Fosterling, Grishenko-Roze, Kouznetsova and colleagues (1996) (as cited in Wasserman, 2003) found that the convergence of the NPQ with verbal self-report indices of the PRF, when translated across multiple languages, tends to be approximately at or above $r = 0.50$. Correlations between the FF-NPQ scales and the NEO-FFI scales appear to be good, with an average correlation of $r = 0.52$, with a self-report and peer rating mean correlation of 0.41, and with an average multiple correlation across 14 external behavior criteria (e.g., number of cigarettes smoked per day) of 0.25 (Paunonen et al., 2001). The normative sample for the NPQ is based on the responses of 1267 individuals from ten different countries. The FF-NPQ norms are based on a sample of 319 North American respondents.

Although both the NPQ and FF-NPQ appear to have adequate psychometric properties that make them sufficient alternatives to the traditional verbal personality measures, according to Wasserman (2003), there are shortcomings. With regard to test construction, some of the items are potentially culture specific. For instance, people from Western cultures tend to value higher education, and therefore such individuals are more likely to endorse an item portraying a character daydreaming about graduation from a university. Such an item may not exemplify familiar behaviors to respondents from non-Western cultures (Wasserman, 2003). Furthermore, although the pictorial stimuli are novel, the quality of the pictures is poor (i.e., they are stick figures). A limitation is the lack of a nationally representative norm group. Finally, knowledge of one's personality profile according to the Big Five personality factors and Murray's (1938) need-based traits is of little value if the goal is to evaluate known or suspected psychopathology as is the case of psychologists needing to establish a diagnosis and treatment plan. Therefore, there is a need for a nonverbal alternative to these nonverbal measures and to the traditional verbally-laden personality measures.

2. STATEMENT OF THE PROBLEM

Psychometrically sound nonverbal personality instruments are needed for fair and accurate psychological evaluations of children and adolescents who have hearing, speech, language, or learning disabilities, who are from different cultural or language backgrounds, and for those who are not verbally communicative due to psychiatric disorders (i.e., selective mutism). Personality tests are necessary tools for mental health practitioners who evaluate the psychological functioning of youth referred for suspected behavioral or emotional problems. In particular, school psychologists are required to evaluate students' psychological strengths and weaknesses when determining if an examinee is "emotionally disturbed." According to Flanagan (1995), "Given the impact that labeling a child emotionally/behaviorally disturbed has on a child's schooling, it behooves school psychologists to use psychometrically defensible methods to make these determinations." (p. 177). Unfortunately, there are no standardized psychometrically sound nonverbal personality measures available, which may account in part for under-identified emotionally disturbed youth who are going without special services. Although the NPQ and FF-NPQ provide good examples of how to measure personality constructs nonverbally, the information obtained is limited and not based on students from the United States. A better instrument is needed.

The conceptual model for the Universal Nonverbal Inventory of Personality (UNIP) is similar to that of the BASC-II (Reynolds & Kamphaus, 2004), Minnesota Multiphasic Personality Inventory-Adolescents (MMPI-A; Butcher, Williams, Graham, Archer, Tellegen, Ben-Porath, & Kaemmer, 1992), Personality Inventory for Children

(PIC; Wirt, Lachar, Klinedinst, Seat, & Broen, 2001), and Clinical Assessment of Behavior (CAB; Bracken and Keith, 2004), which are among the most commonly used clinical and diagnostic personality measures. Unlike the NPQ and FF-NPQ that measure personality characteristics and adjustment in relatively normal individuals, the UNIP is designed to identify psychopathological maladjustment of children who cannot use language to communicate or have limited English language proficiency. The UNIP is designed for ages 8 to 18 years. It provides scales measuring anxiety, depression, atypicality, Attention Deficit Hyperactivity Disorder, social maladjustments, conduct problems, and academic problems, all relevant for evaluating psychological problems in children and adolescents. The nonverbal directions used to administer the UNIP model those used to administer the UNIT (Bracken & McCallum, 1998), which is a nationally standardized assessment of intelligence with entirely nonverbal administration and response formats. The nonverbal directions used with the UNIT have been validated on individuals ranging from 5 to 17 years of age and are the basis for the nonverbal directions that are used to administer the UNIP. Although the UNIP is a promising instrument, designed to overcome the major limitations associated with conventional verbally-laden personality inventories and those presented by the current nonverbal measures of personality, there are no data yet available to support its psychometric integrity. Consequently, the purpose of this research is to evaluate UNIP psychometric properties, including reliability and validity.

3. RESEARCH QUESTIONS

There is a growing need to develop a psychometrically sound nonverbal measure that can be used to evaluate the psychological functioning of children and adolescents who cannot be assessed via traditional language-loaded measurements. General and specific research questions that guide the research are:

1. Does the Universal Nonverbal Inventory of Personality (UNIP), developed by McCallum (in preparation) for this study, provide a psychometrically sound evaluation of child and adolescent psychological functioning? Specific questions include: (a) Are the reliabilities of the seven scales of the UNIP (Anxiety, Depression, Atypicality, Attention Deficit Hyperactivity Disorder, Social Maladjustment, Conduct Problems, and Academic Problems) acceptable ($\geq .70$) as determined by internal consistency via Cronbach's Alpha? (b) Is the reliability acceptable ($\geq .70$) for the Consistency scale?

2. To what extent does the UNIP correlate with a standardized language-based self-report personality measure, namely the BASC-II SRP-A (Reynolds & Kamphaus, 2004)? Specifically, do significant correlations exist between the UNIP Anxiety scale and the BASC-II Anxiety scale, the UNIP Depression scale and the BASC-II Depression scale, the UNIP Atypicality scale and the BASC-II Atypicality scale, the UNIP Attention Deficit Hyperactivity Disorder scale and the BASC-II Hyperactivity scale, the UNIP Attention Deficit Hyperactivity Disorder scale and the BASC-II Attention Problems scale, the UNIP Social Maladjustment scale and the BASC-II Interpersonal Relations scale, the UNIP Academic Problems scale and the BASC-II Attitude to School scale, and the UNIP Academic Problems scale and the BASC-II Attitude to Teachers scale?

4. METHODS

Participants

One hundred and thirty participants completed the UNIP. Responses from these participants were used to determine reliability estimates of the UNIP scales. All but 2 of the participants provided demographic information (e.g., gender, age, date of birth, and academic classification). Participants used for this aspect of study included 45 undergraduate students, 73 graduate students, and 10 participants who classified themselves as "other." There were 96 females and 32 males. Ages of the participants ranged from 19 to 57 years of age, with a mean age of 28 years and standard deviation of 9 years. One hundred of these participants completed the BASC-II SRP-A in addition to the UNIP. Responses from this subset of participants were used to obtain concurrent validity data. Excluding the 2 participants who failed to provide demographic information, this group of participants consisted of 38 undergraduate students, 52 graduate students, and 8 participants who classified themselves as "other." There were 71 females and 27 males. Ages of these participants ranged from 20 to 57 years of age, with a mean age of 28 years and standard deviation of 9 years.

Instruments

McCallum (in preparation) developed the Universal Nonverbal Inventory of Personality (UNIP) for this study (see Appendix A). The UNIP is a self-report, paper and pencil inventory that measures children's and adolescents' responses to emotionally charged situations and is assumed to indirectly assess emotions, thoughts, attitudes, and behavioral dispositions. It can be administered to a group or individually. Items for the

UNIP were created after a review of existing personality literature, nonverbal personality inventories (i.e., NPQ and FF-NPQ) and established verbal personality measures (i.e., BASC-II and MMPI-A). An artist was commissioned to draw illustrations of a gender-neutral and age-neutral central character, drawn to be roughly elementary to middle school age. The drawings were intended to depict the character performing specific behaviors, contemplating specific thoughts, and manifesting specific emotions relevant to the scales of the inventory (see Appendix B).

The original UNIP used in data collection contained 132 items to which participants respond by choosing from two response options: like me and not like me. The response options are represented by two illustrations of facial expressions (i.e., smile and frown) rather than words to avoid the complications associated with written response scales. The UNIP contains eight scales: Anxiety, Depression, Atypicality, Attention Deficit Hyperactivity Disorder, Social Maladjustment, Conduct Problems, and Academic Problems. One validity scale was constructed within the UNIP in an effort to detect invalid responses. Specifically, the Consistency (C) scale was designed to assess examinees' ability to respond conscientiously (i.e., in the same manner to two presentations of the same item). Initially, items were assigned to scales based on theoretical fit.

The UNIP was administered to each of the 130 participants in this study. The BASC-II SRP-A was administered to 100 of those participants. The SRP-A is designed for individuals ages 12 to 21 years. The SRP inventories provide insight into an individual's feelings, thoughts, and emotions by means of 176 true-false questions. The

SRP-A includes the following scales: Anxiety, Attention Problems, Attitude to School, Attitude to Teachers, Atypicality, Depression, Hyperactivity, Interpersonal Relations, Locus of Control, Relations with Parents, and Self-Esteem, Self-Reliance, Sensation Seeking, Sense of Inadequacy, Social Stress, and Somatization. Reliabilities of the BASC-II scales are high, with median internal consistency reliabilities near .80 (Reynolds & Kamphaus, 2004b). The Attitude to School, Atypicality, Social Stress, Anxiety, and Depression are the most reliable scales, with alphas in the middle to upper .80s. The remaining BASC-II scales have reliability alphas generally in the middle .70s to lower .80s, with somewhat lower values for the Self-Reliance and Somatization scales (Reynolds & Kamphaus, 2004b). As reported in Reynolds & Kamphaus (2004b) the validity of the BASC-II scales is supported by several studies in which the SRP-A has been correlated with other self-report measures. For example, correlations between comparable BASC-II SRP-A and Achenbach System of Empirically Based Assessment Youth Self-Report Form (ASEBA; Achenbach & Rescorla, 2001) scales range from .65 to .86. The BASC-II SRP-A and Conners-Wells' Adolescent Self-Report Scale (CASS; Conners, 1997) shows correlations between .52 and .67. The BASC-II SRP-A Depression scale significantly correlates with the Children's Depression Inventory (CDI; Kovacs, 1992) CDI Total score (69). Also, the BASC-II SRP-A Anxiety scale and Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 2000) Total Anxiety scale significantly correlate (.60).

The investigator chose the BASC-II over other self-report measures because it appears to have been developed for the same purposes as those that inspired the

development of the UNIP. For instance, Flanagan (1995) reported that it is the first instrument that aids in the "determination of emotional/behavioral disturbance in accordance with the Individuals with Disabilities Education Act (IDEA)" (p. 178). Similar to the intentions of the UNIP, the BASC was developed with the "purposes of description, educational disability determination (including seriously emotionally disturbed), and DSM-IIIR (American Psychiatric Association, 1987) diagnosis, and to facilitate treatment planning" (Flanagan, 1995, p. 178). It should be noted that although the UNIP and BASC-II SRP-A are designed for children and adolescents, this study was conducted on adults in order to prevent any unintended emotional distress that could be produced in children during the initial phase of development (e.g., some of the pictures depict aggressive acts, which may frighten some children).

Procedures

Students enrolled in selected educational psychology classes located at a southeastern university were allowed to participate in the study during their regularly scheduled class times for extra credit points. Five classes were group-administered the UNIP and the BASC-II SRP-A in counterbalanced order by class and according to guidelines governing participation of human participants. Another class was administered the UNIP only. The sessions began after consent was obtained. Students were asked to record demographic information (e.g., gender, age, and academic classification). The specific directions for the UNIP are listed in Appendix C.

After the students completed the inventories, all materials were collected. Students were informed that the pictures included in the UNIP are not of real people or actual

events. Students were provided with contact information so they could contact the primary researchers and crisis hotlines located in the area in case they experienced emotional distress as a result of the UNIP. Students were not given feedback regarding the results of the UNIP and BASC-II. It should be noted that approximately 15 participants were administered the UNIP and BASC-II on an individual basis as opposed to group-format. The same procedures as described above were followed with those participants. The total time for participation was approximately 55 minutes for the participants who completed the UNIP and the BASC-II SRP-A and 25 minutes for the participants who completed only the UNIP.

Data Analysis

In order to determine the extent to which the Universal Nonverbal Inventory of Personality (UNIP) is a psychometrically sound instrument, the following data analysis techniques were employed.

1. (a) Chronbach's alphas were calculated to determine if the seven scales of the UNIP (Anxiety, Depression, Atypicality, Attention Deficit Hyperactivity Disorder, Social Maladjustment, Conduct Problems, and Academic Problems) display acceptable reliabilities for personality assessment instruments (i.e., $\geq .70$). (b) To determine if the Consistency scale displays acceptable reliability, Pearson product correlations for each of the pairs of Consistency items were computed. Fisher z transformations of those correlations were calculated and then averaged. The Fisher's z average was then transformed to a correlation coefficient via the hyperbolic tangent activation (tanh) function.

2. (a) Pearson product correlations were calculated to determine the extent to which scales devised to assess similar constructs on the UNIP and the BASC-II correlate. Specifically, Pearson product correlation coefficients were calculated to determine the extent to which correlations exist between the UNIP Anxiety scale and the BASC-II Anxiety scale, the UNIP Depression scale and the BASC-II Depression scale, the UNIP Atypicality scale and the BASC-II Atypicality scale, the UNIP Attention Deficit Hyperactivity Disorder scale and the BASC-II Hyperactivity scale, the UNIP Attention Deficit Hyperactivity Disorder scale and the BASC-II Attention Problems scale, the UNIP Social Maladjustment scale and the BASC-II Interpersonal Relations scale, the UNIP Academic Problems scale and the BASC-II Attitude to School scale, and the UNIP Academic Problems scale and the BASC-II Attitude to Teachers scale.

5. RESULTS

Data were obtained to evaluate the basic psychometric properties of the UNIP, including its reliability and concurrent validity. Initially, items were assigned to scales based on theoretical criteria (e.g., items that appeared to depict depressed mood were assigned to the Depression scale). A second round of item assignment to scales was conducted based on item-total scale correlations and the extent to which items contributed to scale reliability. In general, items were assigned to scales if their item-total scale correlation coefficient exceeded .20 and inclusion increased rather than decreased the total scale correlations. Some items considered theoretically appropriate for a scale were retained regardless of their effect on the alpha value. Items that contained either obvious culture-specific content (e.g., graduating from school) or violent content (e.g., shooting a person) were eliminated. Based on these criteria, 43 of the original 132 items were eliminated, leaving a total of 89 items. After revision the Anxiety, Depression, Atypicality, Academic Problems, Attention Deficit Hyperactivity Disorder, and Conduct Problems scales demonstrate adequate reliability; indices for the Social Maladjustment scale and the Consistency scale show marginal reliability. Concurrent validity data are encouraging and provide support for the validity of the UNIP Depression, Anxiety, and Attention Deficit Hyperactivity Disorder scales. In addition, the UNIP Academic Problems scale correlates significantly with the BASC-II Attitude to Teachers scale but does not correlate with the BASC-II Attitude to School scale. Although not a part of the initial research questions, the UNIP Conduct Problems scale significantly correlates with the BASC-II Interpersonal Relations scale. Results provide only marginal support for the

concurrent validity of the UNIP Social Maladjustment scale when correlated with the BASC-II Interpersonal Relations scale and the UNIP Atypicality scale when correlated with the BASC-II Atypicality scale. Descriptive statistics for each of the UNIP and BASC-II scales, including means and standard deviations, are provided in Tables D1* and D2. The BASC-II mean scores were slightly below "average" in general but relatively close to the mean from the general population (i.e., mean of 50 and standard deviation of 10). More detailed analyses follow.

Reliability of the UNIP

Reliability was measured by Cronbach's alpha, a measure of internal consistency. DeVellis (1991) recommends that an alpha below .60 be considered unacceptable; between .60 and .65 undesirable; between .65 and .70 minimally acceptable; between .70 and .80 acceptable; between .80 and .90 very good; and above .90 excellent.

Cronbach's alphas can be found in Table D3. The Anxiety and Depression scales show the strongest reliability indices with alphas of .85 and .84 respectively. The Atypicality scale (.78), Academic Problems scale (.77), Attention Deficit Hyperactivity Disorder scale (.75), and Conduct Problems scale (.73) each produce acceptable reliabilities. The Social Maladjustment scale produces the poorest reliability with a Cronbach's alpha of .64, a value suggesting that the scale needs to be improved.

The Consistency scale shows relatively poor reliability (i.e., the correlation coefficient is .54). Another strategy to evaluate the utility of the consistency scale is to determine how sensitive it is to random responding. As can be seen in Table D4, 37% of

* All tables are located in Appendix.

the sample completed all the consistency items appropriately. An additional thirty-five percent had only 1 inconsistent response and 20% had only 2 inconsistent responses, leaving only 8% with 3 or more inconsistent responses, indicating strong sensitivity to random responding for this sample.

Item-total correlations for each scale are provided in Table D5. Compared to the other UNIP scales, the Anxiety scale has the widest range of item-total correlations with a range of .18 to .66 and mean correlation of .41. The item-total correlations for the Depression scale range from .23 to .62 with a mean correlation of .41 as well. For the Atypicality scale, item-total correlations range from .23 to .59 with a mean correlation of .43. Item-total correlations for the Academic Problems scale range from .31 to .65 with a mean correlation of .51, which represents the strongest mean correlation among the UNIP scales. For the Attention Deficit Hyperactivity Disorder scale, item-total correlations range from .21 to .51 with a mean correlation of .34. The Conduct Problems scale has item-total correlations that range from .29 to .60 with a mean correlation of .41. The Social Maladjustment scale has the weakest range of item-total correlations and lowest mean correlation. The item-total correlations for the Social Maladjustment scale range from .21 to .38 with a mean correlation of .29. Overall, the UNIP item-total correlations range from .18 to .66 with a mean correlation of .39.

Validity of the UNIP

Pearson product correlations were obtained to determine the extent to which theoretically comparable scales from the UNIP and BASC-II SRP-A correlate. Table D6 contains the Pearson product correlations between UNIP scales and BASC SRP-A scales

thought to measure similar constructs. Examination of the correlation coefficients indicates that the UNIP Depression scale and BASC-II Depression scale are significantly and strongly related ($r = .62, p < .01$). Of interest, the UNIP Anxiety scale and the BASC-II Depression scale shows a strong correlation coefficient ($r = .57, p < .01$). The UNIP Attention Deficit Hyperactivity Disorder scale significantly correlates with the BASC-II Attention Problems scale ($r = .40, p < .01$). Although not as strong as the correlation between the UNIP Anxiety scale and the BASC-II Depression scale, a significant correlation exists between the UNIP Anxiety scale and the BASC-II Anxiety scale ($r = .37, p < .01$). The UNIP Academic Problems scale significantly correlates with the BASC-II Attitude to Teachers scale ($r = .36, p < .01$). The UNIP Attention Deficit Hyperactivity Disorder scale and the BASC-II Hyperactivity scale significantly correlates ($r = .34, p < .01$). The UNIP Conduct Problems scale significantly correlates with the BASC-II Interpersonal Relations scale ($r = .32, p < .01$). A significant but low correlation exists between the UNIP Social Maladjustment scale and the BASC-II Interpersonal Relations scale ($r = .27, p < .01$). The UNIP Academic Problems scale does not significantly correlate with the BASC-II Attitude to School scale ($r = .17$). Results indicate that the UNIP Atypicality scale and the BASC-II Atypicality do not significantly correlate ($r = .04$).

6. DISCUSSION

The UNIP is a newly developed personality test designed to identify emotional, social, and academic impairment in children and adolescents who cannot use language to communicate. A personality test of this nature is needed for youth who have hearing, speech, language, or learning disabilities, who are from different cultural or language backgrounds, and for those who are verbally uncommunicative due to psychiatric disorders (i.e., selective mutism). Since there are no data yet available to support its psychometric integrity, this study was conducted to address that limitation. Reliability and concurrent validity data are discussed and guide recommendations for improving the UNIP. Also, a comparison of the reliability coefficients of the BASC-II scales and the UNIP scale was conducted as a basis of evaluating the UNIP. The BASC-II reliability coefficients used for this comparison are from the BASC-II SRP-A general norm sample that represents adolescents between the ages of 15 and 18 years (Reynolds & Kamphaus, 2004b).

Internal Consistency

Reliability data for the Anxiety scale are encouraging. Compared to the other UNIP scales, the Anxiety scale has the strongest reliability with a Cronbach's alpha of .85, which is comparable to the BASC-II Anxiety scale alpha of .86. The Anxiety items depict the central character engaging in behaviors that are typically associated with anxiety, worry, and nervousness (e.g., biting one's nails and shaking). The context of the Anxiety items range from worrying about making bad grades, getting in trouble with a

teacher or parents, and having accidents (e.g., dropping lunch tray) to having nightmares, worrying about sickness and injuries, and feeling anxious in crowds.

The Depression scale yielded a Cronbach's alpha of .84, which is slightly lower than the BASC-II Depression scale alpha of .86. The UNIP Depression items depict typical symptoms and behaviors associated with depression such as crying, suicidal ideation, withdrawal, and self-depreciation. Five items were found to correlate with both the Anxiety and Depression scales. This overlap is understandable given that anxiety and depression are assumed to have similar emotional features. For instance, as cited in Blumberg and Izard (1986), Bartlett and Izard (1972) compared the emotion patterns of anxious and depressed adults. Anxious adults reported patterns of emotions in which fear was the central emotion with interest, anger, guilt, shame, and shyness as variable emotions in the pattern. The depressed adults reported sadness as the central emotion and anger, guilt, self-direction, hostility, shame, and fear as variable components in the pattern. It should be noted that the Anxiety and Depression scales have the highest numbers of items with 23 and 22 items respectively. Although item number alone is not sufficient to account for the higher reliability alphas, it is possible that the large number of items within the scales contributed to relatively higher scores.

Reliability data for the Atypicality scale suggest acceptable reliability with a Cronbach's alpha of .78. This alpha is slightly lower than the BASC-II Atypicality scale alpha of .82. The UNIP Atypicality items depict relatively bizarre behaviors and situations such as jumping over a fully grown tree, flying amongst the clouds, talking with animals, seeing monsters, flying a rocket, climbing a skyscraper, and holding a

piano above one's head. Two of the items are intended to measure respondents' tendency to experience visual and/or auditory hallucinations. The uniqueness of the 11 UNIP Atypicality items is supported by the fact that none of the items correlate with any of the other scales.

The Academic Problems scale demonstrates an acceptable reliability level with a Cronbach's alpha of .77, which is slightly lower than the BASC-II Attitude to School scale alpha of .82 and BASC-II Attitude to Teachers scale alpha of .79. It should be noted that the Academic Problems scale contains the lowest number of items with only 8 items.

The Attention Deficit Hyperactivity Disorder scale displays acceptable reliability with a Cronbach's alpha of .75. This value is slightly higher than the BASC-II Hyperactivity scale alpha of .74 and slightly lower than the BASC-II Attention Problems scale alpha of .79. The Attention Deficit Hyperactivity Disorder scale consists of items designed to depict inattentive, hyperactive, and impulsive behaviors. The central character is illustrated daydreaming during class, not paying attention to tasks (e.g., reading a book while riding a bicycle), being unorganized (e.g., messy room), being overly active compared to other children, and impulsive (e.g., throwing a brick through a window). The Attention Deficit Hyperactivity Disorder scale has approximately 8 items designed to measure hyperactivity and impulsivity and 9 items designed to measure inattentiveness for a total of 17 items. Initially, like the BASC-II, those behaviors were intended to be measured separately. The original UNIP contained an Attention Problems scale and separate Hyperactivity scale. The reliabilities of those separate scales were

unacceptably low. As a result, the scales were combined and revised, resulting in the current Attention Deficit Hyperactivity Disorder scale.

The Conduct Problems scale displays acceptable reliability with a Cronbach's alpha of .73. Of interest, this value is slightly lower than the BASC-II Interpersonal Relations scale alpha of .78. Although the Conduct Problems scale measures a few of the same impulsive behaviors as measured by the Attention Deficit Hyperactivity Disorder, it primarily focuses on rule-breaking and antisocial behaviors. For this scale the central character is illustrated yelling at another child, dropping a cat off of a building, burning a book, setting fire to a tree, eating messily, fighting, kicking a cat, and stealing. The Conduct Problems scale consists of 10 items that represent clear examples of inconsiderate and disruptive behaviors.

The Social Maladjustment scale displays the poorest reliability with a Cronbach's alpha of .64, which is considered undesirable and indicates that the scale needs to be improved. The reliability of the Social Maladjustment scale is somewhat lower than that of the BASC-II Interpersonal Relations scale alpha of .78. The scale contains 13 items with item-total correlations reaching only .38. The scale consists of items that are intended to depict the central character being selfish (e.g., not sharing toys), showing poor social skills (e.g., yelling at and teasing others), and not participating in group activities (e.g., playing while the other children are cleaning). It is possible that the scale failed to produce a high reliability index because it is unclear in many of the items if the central character is engaging in the maladaptive behaviors intentionally (e.g., being defiant or disrespectful), which is better measured by the Conduct Problems scale, or

unintentionally due to a lack of social skills. In an attempt to address the reliability of the Social Maladjustment scale, an effort was made to combine the Social Maladjustment scale with the Conduct Problems scale. However, combining the scales failed to improve the reliability of the Social Maladjustment scale and the effort was discontinued. Also, many of the items intended to measure social problems did not correlate highly with the total score but correlated more highly with the depression scale score and were deleted from the Social Maladjustment scale.

The Consistency scale consists of 9 pairs of duplicated items that are scattered throughout the UNIP. To illustrate, item 1 is identical to item 108, and therefore represents a Consistency item pair. The Consistency scale is intended to measure the extent to which respondents are careless in responding. Results indicate that the Consistency scale is only minimally reliable with a correlation of .54. This level of reliability falls well below the desired level ($\geq$.70) and indicates that the scale needs to be revised. The UNIP Consistency scale was intended to serve a similar purpose as that of the BASC-II Consistency Index, which detects when an individual's responses are not internally consistent (i.e., respondent answers differently to very similar items).

Overall, the reliability of the UNIP scales range from very good to undesirable. The Anxiety and Depression scales show the strongest reliabilities indicating that the items within those scales are more homogeneous and therefore are measuring the same underlying property. The Atypicality, Attention Deficit Hyperactivity Disorder, Conduct Problems, and Academic Problems scales demonstrate acceptable reliabilities. Each of the UNIP scales, with the exception of the Social Maladjustment scale, compared

favorably to similar BASC-II scales. These favorable comparisons are particularly encouraging considering the UNIP's relatively small sample size, which would likely weaken scale reliabilities. The Social Maladjustment scale has the weakest reliability. The Consistency scale shows unacceptable reliability and should be revised.

Concurrent Validity of the UNIP with the BASC-II

The purpose of the second research question was to evaluate the concurrent validity of the UNIP by correlating UNIP and BASC-II SRP-A scales that are thought to measure similar constructs. Each of the UNIP scales were developed using comparable BASC-II scales as points of reference, and one would expect that the related scales would correlate with each other. Some of the hypothesized relationships between comparable UNIP and BASC-II scales occur.

Results indicate that the UNIP Depression scale and the BASC-II Depression scale show the strongest correlation with a validity coefficient of .62 ($p < .01$). As hypothesized, the UNIP Anxiety scale and BASC-II Anxiety scale show a significant correlation with a validity coefficient of .37 ($p < .01$). In an attempt to better understand what the UNIP Anxiety scale measures, the researcher correlated the UNIP Anxiety scale with other scales (e.g., BASC-II Depression scale). This analysis is logical given that anxiety is a commonly associated symptom of other problems such as depression among children and adolescents (House, 2002). Results of this analysis indicate that the UNIP Anxiety scale is strongly related to the BASC-II Depression scale ($r = .57$, $p < .01$), which is not surprising given the comorbidity literature: Brady and Kendall (1992) found that 15.9% to 61.9% of children and adolescents identified as anxious or depressed have

comorbid anxiety and depression disorders. These findings support the idea that the UNIP Anxiety and the UNIP Depression scales are valid measures of internalized emotional difficulties that are typically comorbid.

Although the correlations are only modest, the UNIP Attention Deficit Hyperactivity Disorder scale correlates significantly with both the BASC-II Hyperactivity scale ($r = .34$, $p < .01$) and the BASC-II Attention Problems scale ($r = .40$, $p < .01$). These findings provide limited initial support for the UNIP Attention Deficit Hyperactivity Disorder scale, which is encouraging considering the difficulty associated in creating clear depictions of attention problems. The scale contains approximately the same number of items devised to measure impulsivity, hyperactivity, and inattention problems. Thus, both constructs are considered. Overall, the UNIP Attention Deficit Hyperactivity Disorder scale appears to be a reliable and valid measure of Attention Deficit Hyperactivity Disorder (ADHD) symptoms. The scale could be helpful in screening children and adolescents who are "at-risk" for an ADHD diagnosis. Elevated scores on this scale coupled with home and school behavior ratings would provide clinicians with valuable information for diagnosis.

Results are mixed regarding the validity of the UNIP Academic Problems scale. Data indicate that the UNIP Academic Problems scale significantly correlates with the BASC-II Attitude to Teachers scale ($r = .36$, $p < .01$). However, the UNIP Academic Problems scale does not correlate significantly with the BASC-II Attitude to School Scale ($r = .17$, $p > .05$). These findings are logical considering a majority of the UNIP Academic Problems items portray the central character having either positive or negative

interactions with a female authority figure that could easily be interpreted as a teacher. Also, the BASC-II Attitude to School scale focuses more on an individual's general opinion of the utility of school, whereas the BASC-II Attitude to Teachers scale assesses an individual's perception of teachers as being unfair, uncaring, or unmotivated to assist the student (Reynolds & Kamphaus, 2004b). Based on the results, the UNIP Academic Problems scale appears to be a more narrow measure of student-teacher relationships as opposed to a broad measure of school functioning. This type of measure is valuable since an elevated score on this scale would indicate the presence of personality conflicts with teachers that should be addressed with an intervention.

Following revisions of the UNIP Conduct Problems scale, the researcher correlated the UNIP Conduct Problems scale with the BASC-II Interpersonal Relations scale. Four of the 10 final Conduct Problems items depict antisocial behaviors (e.g., yelling at another person, stealing from another person, fighting, and eating messily) that likely interfere with positive relations with others. Results indicate that the UNIP Conduct Problems scale significantly correlates with the BASC-II Interpersonal Relations scale ($r = .32$, $p < .01$). Although the UNIP Conduct Problems scale is likely a good indicator of rule-breaking tendencies, current findings suggest that it may also tap perceived ability to relate to others and to develop social skills.

It is not surprising that results indicate there is little correlation between the UNIP Social Maladjustment scale and the BASC-II Interpersonal Relations scale ($r = .27$, $p < .01$) given the poor reliability of that UNIP scale. As discussed previously it was difficult to create items that clearly depict the presence of social skill deficits and lack of

friendships as opposed to items that depict more serious and intentional acts of defiance, selfishness, and aggression. Overall, results provide only limited support for the validity of the UNIP Social Maladjustment scale.

Results indicate a lack of correlation between the UNIP Atypicality scale and BASC-II Atypicality scale ($r = .04$, $p > .05$). One possible explanation for this finding is that it is difficult to design items that measure paranoia, which is one focus of the BASC-II Atypicality items. Initially the UNIP Atypicality scale contained items aimed at measuring respondents' paranoia (e.g., depictions of others hurting the examinee). The items vividly depicted the central character being hurt by another person (e.g., girl holding a knife) and the central character hurting another person (e.g., shooting a boy). However, due to the graphic nature of the items, they were deleted from the scale. Also, two UNIP items measure the presence of visual and auditory hallucinations. More items of this nature were created, but the complexity of the pictures made them difficult to understand. Those pictures were dropped from the UNIP as well. Overall, results indicate that the UNIP Atypicality scale reliably measures a construct other than that which is measured by the BASC-II Atypicality scale (i.e., the measure is reliable but not valid when the BASC-II Atypicality scale is used as the standard). The UNIP Atypicality scale most likely assesses emotional affect that could reasonably be labeled as psychosis (i.e., loss of contact with reality, hallucinations, delusions, and problems in thinking clearly).

In an effort to evaluate the sensitivity of the Consistency scale, inconsistent responding between Consistency pairs was tabulated. This analysis was conducted to establish the level at which respondents' answers to items should be interpreted with

caution and considered "at-risk" for careless responding (see Table D4). For some purposes "at-risk" status is considered to exist when one's inconsistency score is equal to or greater than the highest/lowest 16% of a sample. For comparison purposes thirty-seven percent of the sample has no inconsistencies. Only 8% had more than two inconsistent responses. These data suggest that results should be interpreted with caution when individuals respond inconsistently more than twice. It is encouraging that the Consistency scale has such a highly skewed frequency distribution. This type of distribution is desirable for validity indexes if it is assumed that only a small number of cases are invalid and most cases are valid (Reynolds & Kamphaus, 2004b). The BASC-II software program provides a similar measure of inconsistent responding with the Consistency Index. The Consistency Index consists of 20 item pairs that have the highest item-correlations, therefore suggesting that the items should be answered similarly. The Consistency Index is scored by summing the absolute values of the score differences between each set of Consistency Index item pairs. A high score on the Consistency Index indicates that the results should be interpreted cautiously. Based on the normative sample, scores between 17 and 25 suggest caution in interpretation, while scores of 26 or higher suggest extreme caution in interpretation. These cutoffs were designed to identify approximately 5% or less of the cases in which there was random responding. Based on these results, it appears that the UNIP Consistency scale, with some minor revisions aimed at increasing reliability (e.g., more item pairs, clearer items) could serve as a similar measure of validity.

Summary and Implications

Currently there are no commercially available validated nonverbal measures to assess the emotional and behavioral functioning of youth who cannot use language to communicate. This study was conducted to evaluate the psychometric properties of the UNIP, which is designed to overcome the major limitations of traditional verbal personality tests. Results indicate that the UNIP Anxiety scale and the UNIP Depression scales appear to be reliable and valid measures of internalized emotional problems. Results support the reliability and concurrent validity of the UNIP Attention Deficit Hyperactivity Disorder scale, suggesting it could eventually be used as a self-report Attention Deficit Hyperactivity Disorder screening measure. The UNIP Academic Problems scale appears to be a reliable and valid measure of student-teacher relationship issues. The UNIP Conduct Problems scale appears to be a reliable and valid measure of respondents' thoughts regarding their ability to relate to others. Since a majority of the Conduct Problems items focus on rule-breaking and serious antisocial behaviors, it is conceivable that the scale is a broad measure of serious problematic behaviors (e.g., breaking rules, impulse control, bullying) that adversely affect school, community, and home interactions. Based on the lack of reliability and concurrent validity of the UNIP Social Maladjustment scale, it should be either eliminated or strengthened. The UNIP Atypicality scale shows acceptable reliability. Unfortunately it does not appear to measure the same construct that is measured by the BASC-II Atypicality scale. In its current form the Atypicality scale may best be called the psychosis scale.

Limitations and Suggestions for Further Research

One of the most obvious limitations of this study is the nature of the sample. Although the UNIP is ultimately intended for children and adolescents, the researcher chose to conduct this study on adults in order to prevent any unintended emotional distress that could be produced in children during this initial phase of test development. As a result of this decision, the sample of participants in this study came primarily from one southeastern university. Since many of the items on the UNIP portray behaviors and thoughts that are typical of children who experience emotional distress, it is conceivable that the adult participants could not relate to the item content (e.g., having peers reject them on the playground, dropping their food tray in from of peers, not winning in a race). In the future, a study should be conducted with school-aged children to determine the feasibility of using the UNIP with that population. Future research should include clinical comparisons of groups of children with and without emotional problems. These data could allow for differential validity analysis. Furthermore, the sample size in this study is small, especially given the nature of the research questions addressed. It would be beneficial to conduct this research on a larger sample to allow stronger multivariate analyses (e.g., factor analysis).

The concurrent validity indices between the UNIP and the BASC-II may be impacted (and reduced) because of different methods used to evaluate pathology within the instruments. That is, one is nonverbal and one is verbal. Ideally, the concurrent validity of the UNIP would have been evaluated using a similar nonverbal self-report

instrument, thereby minimizing method variance. However, as previously discussed, there currently are no nonverbal measures of psychopathology available.

Since the UNIP Consistency scale allows researcher to identify careless responders, it would be beneficial to eliminate the inconsistent respondents (i.e., those who have 3 or more inconsistent responses) from the sample prior to reliability and validity data analysis. Eliminating the inconsistent responders would likely increase the reliability and validity of the scales.

The value of including both Anxiety and Depression scales should be investigated. According to Brady and Kendall (1992), depressed children tend to score high on both depression and anxiety measures, while anxious children tend to score high on anxiety measures but low on depression measures. Future research should investigate whether this trend is true for the UNIP Depression and Anxiety scales.

As noted by Paunonen et al. (1990), there are many problems associated with the development of a nonverbal personality test. For instance, it is difficult to create understandable and suitable items for scales. This disadvantage is especially true when trying to construct pictorial items intended to measure internal thoughts and feelings. Many of the behaviors, thoughts, and symptoms of emotional and behavioral problems can be described verbally but are increasingly difficult to portray in a nonverbal item thereby limiting the range of symptoms of problems that can be measured. This limitation was exemplified when the UNIP author attempted to construct items for the UNIP Atypicality scale. It was increasingly difficult to create nonviolent yet understandable items to measure paranoia.

For the most part, the results of this study are encouraging and suggest that the UNIP has promise. This study represents the first step in validating the UNIP. The goal of the research is to produce a measure that allows professionals to gain a better understanding of the internal experiences of youth who cannot or do not use language to communicate. Modifications of the current UNIP and further research are needed. Until those revisions are made and further research is conducted, these results should be considered tentative.

LIST OF REFERENCES

LIST OF REFERENCES

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Aiken, L. R. (1999). *Personality assessment: Methods and practices (3rd ed., revised).* Seattle: Hogrefe & Huber Publishers.

Aiken, L. R. (1997). *Assessment of adult personality.* New York: Springer Publishing Company.

Aiken, L. R. (1995). *Assessment of adult personality.* New York: Springer Publishing Company.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd rev. ed.).* Washington, DC: Author.

Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin, 111*(2), 244-255.

Beck, A. T., & Steer, R. A. (1987). *Beck depression inventory manual.* San Antonio, TX: The Psychological Corporation.

Bender, L. (1938). *A visual motor gestalt test and its clinical use. Research monograph No. 3.* New York: American Orthopsychiatric Association.

Blumberg, S. H., & Izard, C. E. (1986). Discriminating patterns of emotions in 10- and 11-year-ol children's anxiety and depression. *Journal of Personality and Social Psychology, 51*, 852-857.

Bracken, B. A., & Keith, L. K. (2004). *Clinical Assessment of Behavior.* Lutz, FL: Psychological Assessment Resources.

Bracken, B.A, & McCallum, R. S. (1998). *Universal Nonverbal Intelligence Test: Examiner's manual.* Itasca, IL: Riverside.

Briggs, K. C., & Myers, I. B. (1985). *Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Butcher, J. N., Dalhstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota multiphasis personality inventory—2 (MMPI-2): Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

Butcher, J. N., Williams, C. L., Graham, J. R., Archer, R. P., Tellegen, A., Ben-Porath, Y. S., & Kaemmer, B. (1992). *MMPI-A (Minnesota Multiphasic Personality Inventory-Adolescent): Manual for administration, scoring, and interpretation.* Minneapolis: University of Minnesota Press.

Cattell, R. B., Eber, H. J., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Champaign, IL: Institute for Personality and Ability Testing.

Center for Effective Collaboration and Practice. (2001). Students with emotional disturbance: Eligibility and characteristics. Retrieved September 10, 2003, from http://www.cecp.air.org/resources/20th/eligchar.htm

Conners, C. K. (1997). *Conners' Rating Scales-Revised.* North Towanda, NY: Multi-Health Systems.

Costa, P. T., Jr., & McCraeR. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual.* Psychological Assessment Resources: Odessa, FL.

Derogatis, L. (1994). *SCL-90-R: Symptom Checklist-90-R. Administration, scoring, and procedural manual.* Minneapolis, MN: National Computer Systems.

DeVellis, R. F. (1991). *Scale development: Theory and applications.* Newbury Park, CA: Sage.

Edwards, A. L. (1959). *Edwards Personal Preference Schedule manual.* New York: The Psychological Corporation.

Flanagan, R. (1995). A review of the Behavior Assessment System for Children (BASC): Assessment consistent with requirements of the Individuals with Disabilities Education Act (IDEA). *Journal of School Psychology, 33*(2), 177-186.

Forness, S. R., & Knitzer, J. (1992). A new proposed definition and terminology to replace "Serious Emotional Disturbance" in Individuals with Disabilities Education Act. *School Psychology Review, 21*, 12-20.

Frisby, C. L. (1999). Straight talk about cognitive assessment and diversity. *School Psychology Quarterly, 14*, 195-207. In R. S. McCallum, B. A. Bracken, & J. D. Wasserman (Eds.), *Essentials of nonverbal assessment.* New York: John Wiley & Sons.

Goh, D. S., Teslow, C. J., & Fuller, G. B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology, 12*, 696-706.

Gough, H. G. (1987). *The California Psychological Inventory administrator's guide.* Palo, Alto, CA: Consulting Psychologists Press.

House, A. E. (2002). *DSM-IV diagnosis in the schools.* New York, NY: Guilford Press.

Jackson, D. N. (1984). *Personality research form manual.* Port Huron, MI: Research Psychologists.

Kauffman, J. M. (1994). *Characteristics of behavior disorders of children and youth (5th ed.).* Columbus, OH: Merrill.

Kidder-Ashley, P., Deni, J. R., Azar, K. R., & Anderton, J. B. (2000). Comparison of 40 states' procedures for identifying students with serious educational problems. *Education, 120* (3), 558-68.

Kindler, A. L. (2002). *Survey of the states' limited English proficient students and available educational programs and services 1999-2000 summary report.* Washington, DC: National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs.

Knoff, H. M. (1983). Personality assessment in the schools: Issues and procedures for school psychologists. *School Psychology Review, 12*(4), 391-398.

Knoff, H. M. (1986). *Assessment of child and adolescent personality.* New York: Guilford Press.

Knoff, H. M. (1995). Best practices in personality assessment. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology-III* (pp. 1281-1301). Bethesda, MD: The National Association of School Psychologists.

Kovacs, M. (1992). *Children's Depression Inventory: Manual.* North Towanda, NY: Multi-Health Systems.

Lachar, D., & LaCombe, J. A. (1983). Objective personality assessment: The personality inventory for children and its applications in the school setting. *School Psychology Review, 12*(4), 399-406.

McCallum, R. S. (2003). *Handbook of Nonverbal Assessment.* New York, NY: Kluwer Academic/Plenum Publishers.

McCallum, R. S., Bracken, B. A., & Wasserman, J. D. (2001). *Essentials of nonverbal assessment*. New York: John Wiley & Sons.

Million, T. (1987). *Manual for the MCMI-II.* Minneapolis, MN: National Computer Systems.

Morgan, C. D., & Murray, H. A. (1935). A method for investigating phantasies: The Thematic Apperception Test. *Archives of Neurology and psychiatry, 34*, 289-306.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs. (2002). *United States rate of LEP growth.* Retrieved September, 11, 2004, from http:www.ncbe.gwu.edu/states/stateposter.pdf

Ochoa, S. H., Riccio, C., Jimenez, S., Alba, R. G., & Sines, M. (2004). Psychological assessment of English language learners and/or bilingual students: An investigation of school psychologist' current practices. *Journal of Psychoeducational Assessment, 22*, 93-105.

Oswald, D. P., & Coutinho, M. J. (1995). Identification and placement of students with serious emotional disturbance. Part I: Correlates of state child-count data. *Journal of Emotional and Behavioral Disorders, 3*(4), 224-229.

Paunonen, S. V., & Ashton, M. C. (2002). Nonverbal assessment of personality: The NPQ and the FF-NPQ. In B. De Raad, & M. Perugini (Eds.), *Big five assessment* (pp. 171-194). Seattle; Toronto: Hogrefe & Huber Publishers.

Paunonen, S. V., Ashton, M. C., & Jackson, D. N. (2001). Nonverbal assessment of the big five personality factors. *European Journal of Personality, 15*, 3-18.

Paunonen, S. V., Ashton, M. C., & Jackson, D. N. (2000). *Five-Factor Nonverbal Personality Questionnaire (FF-NPQ)*. Port Huron, MI: Sigma Assessment Systems.

Paunonen, S. V., & Jackson, D. N. (1998). *Nonverbal Personality Questionnaire (NPQ).* Port Huron, MI: Sigma Assessment Systems.

Paunonen, S. V., Jackson, D. N., & Keinonen, M. (1990). The structured nonverbal assessment of personality. *Journal of Personality, 58*(3), 481-502.

Paunonen, S. V., Keinonen, M., Trzebinski, J., Fosterling, F., Grishenko-Roze, N., Kouznetsova, L, & Chan, D. W. (1996). Structure of personality in six cultures. *Journal of Cross-Cultural Psychology, 27*(3), 339-353.

Prout, H. T. (1983). School psychologists and social-emotional assessment techniques: Patterns in training and use. *School Psychology Review, 12*(4), 377-383.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children, Second Edition (BASC-II).* Circle Pines, MN: American Guidance Service.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children, Second Edition Manual.* Circle Pines, MN: American Guidance Service.

Reynolds, C. R., & Richmond, B. O. (2000). *Revised Children's Manifest Anxiety Scale.* Los Angeles: Western Psychological Services.

Rorschach, H. (1921). *Psychodiagnostik.* Bern: Ernst Bircher.

Segal, D. L., & Coolidge, F. L. (2003). Objective assessment of personality and psychopathology: An overview. In M. J. Hilsenroth, D. L. Segal, & Michel Hersen (Eds.), *Comprehensive handbook of psychological assessment, volume 2 (pp. 3-13).* New York: John Wiley & Sons.

Tennessee State Department of Education. (n.d.). *Special Education Manual 2003.* Retrieved June 7, 2003, from http://www.state.tn.us/education/speced/semanual.pdf

Wasserman, J. D. (2003). Nonverbal assessment of personality and psychopathology. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 283–313). New York: Kluwer Academic/Plenum Publishers.

Watkins, C. E., Campbell, V. L., Neiberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*, 54-60.

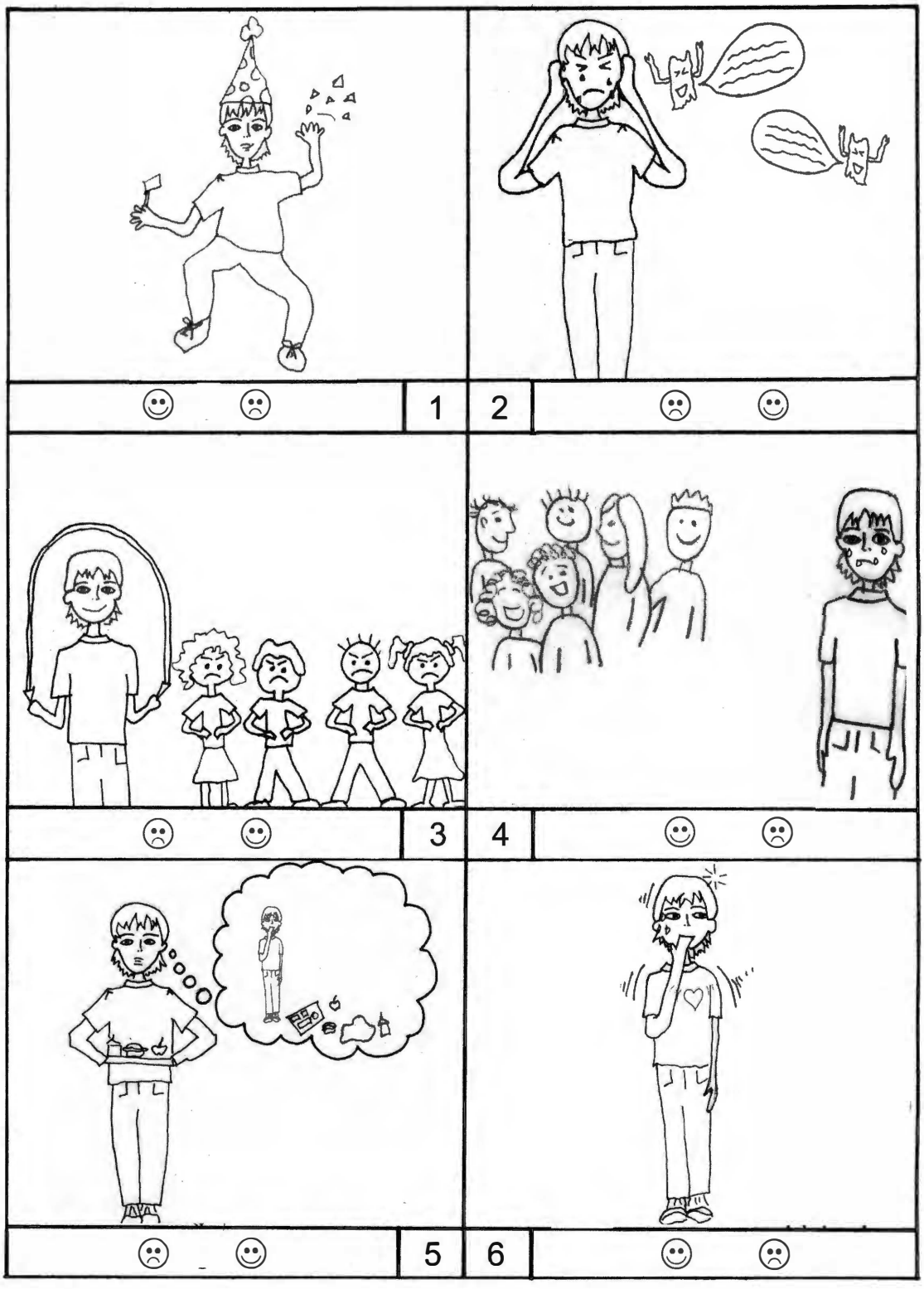
Wirt , R. D., Lachar, D., Klinedinst, J. E., Seat, P. D., & Broen, W. E. (2001). *Multidimensional evaluation of child personality: A manual for the Personality Inventory for Children (2nd ed.).* Los Angeles: Western Psychological Services.

Woodworth, R. S. (1917). *Personal data sheet.* Chicago: C. H. Stoelting Company.

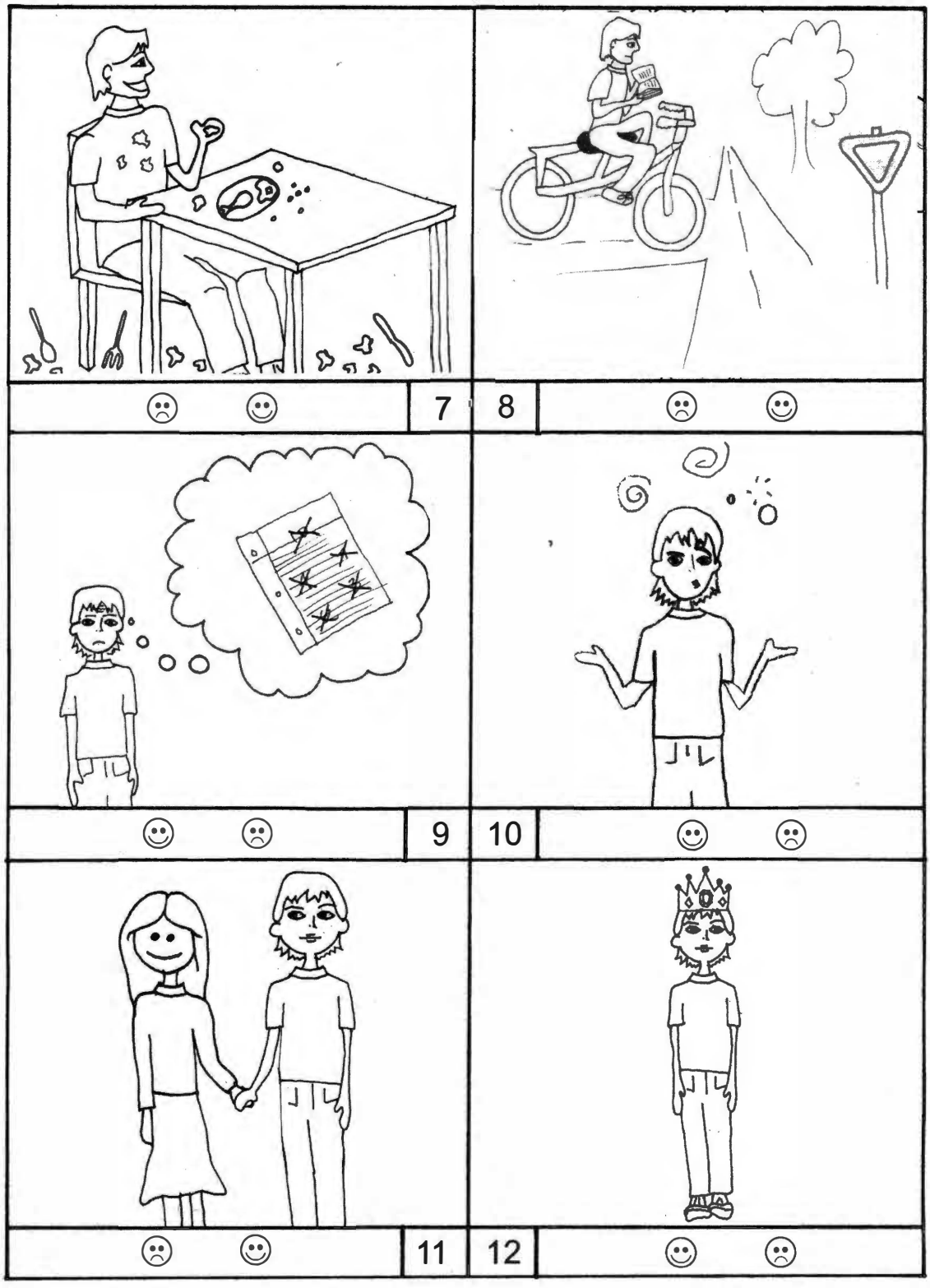
APPENDICES

Appendix A - Universal Nonverbal Inventory of Personality

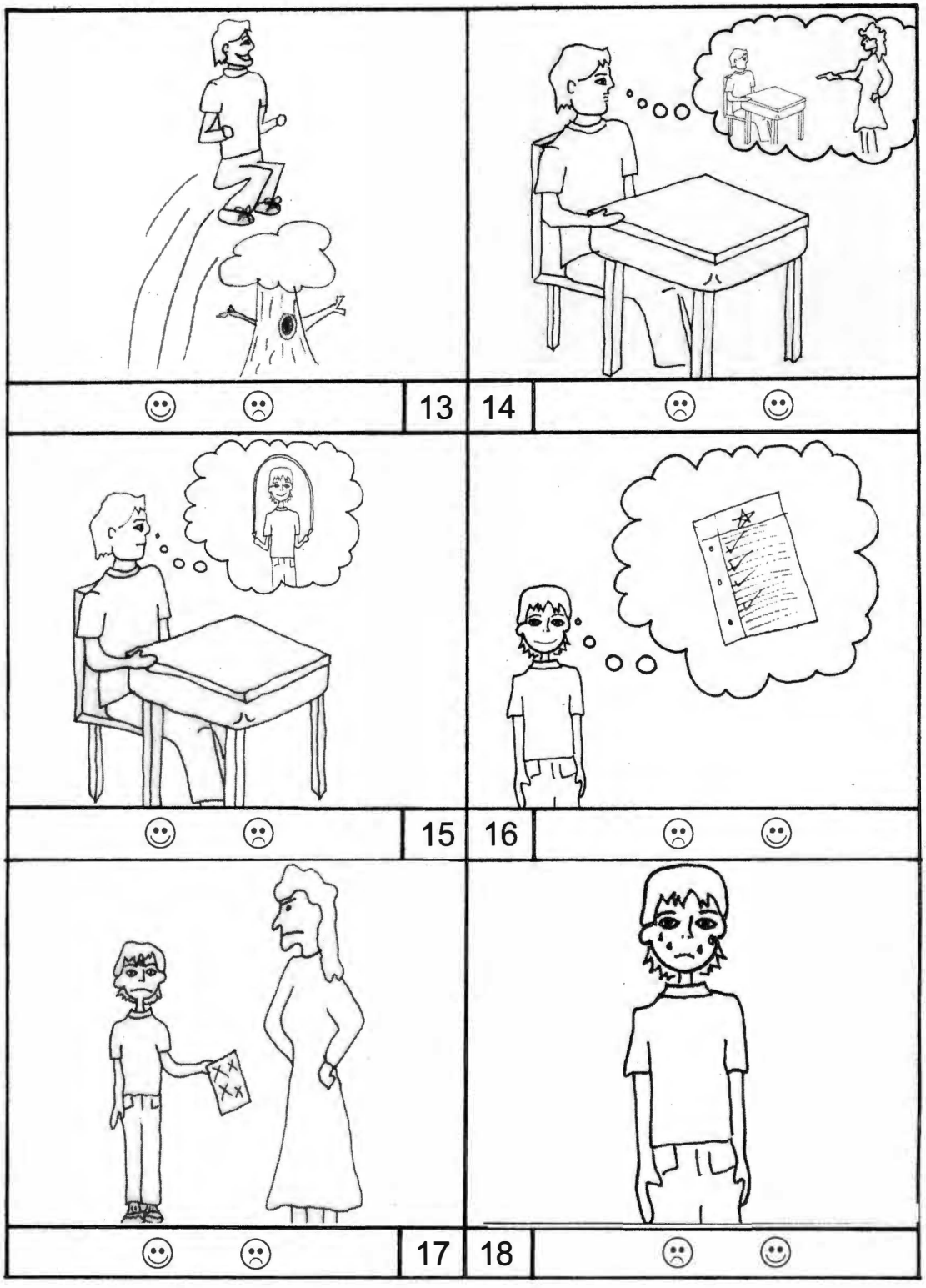
Appendix A. Continued.



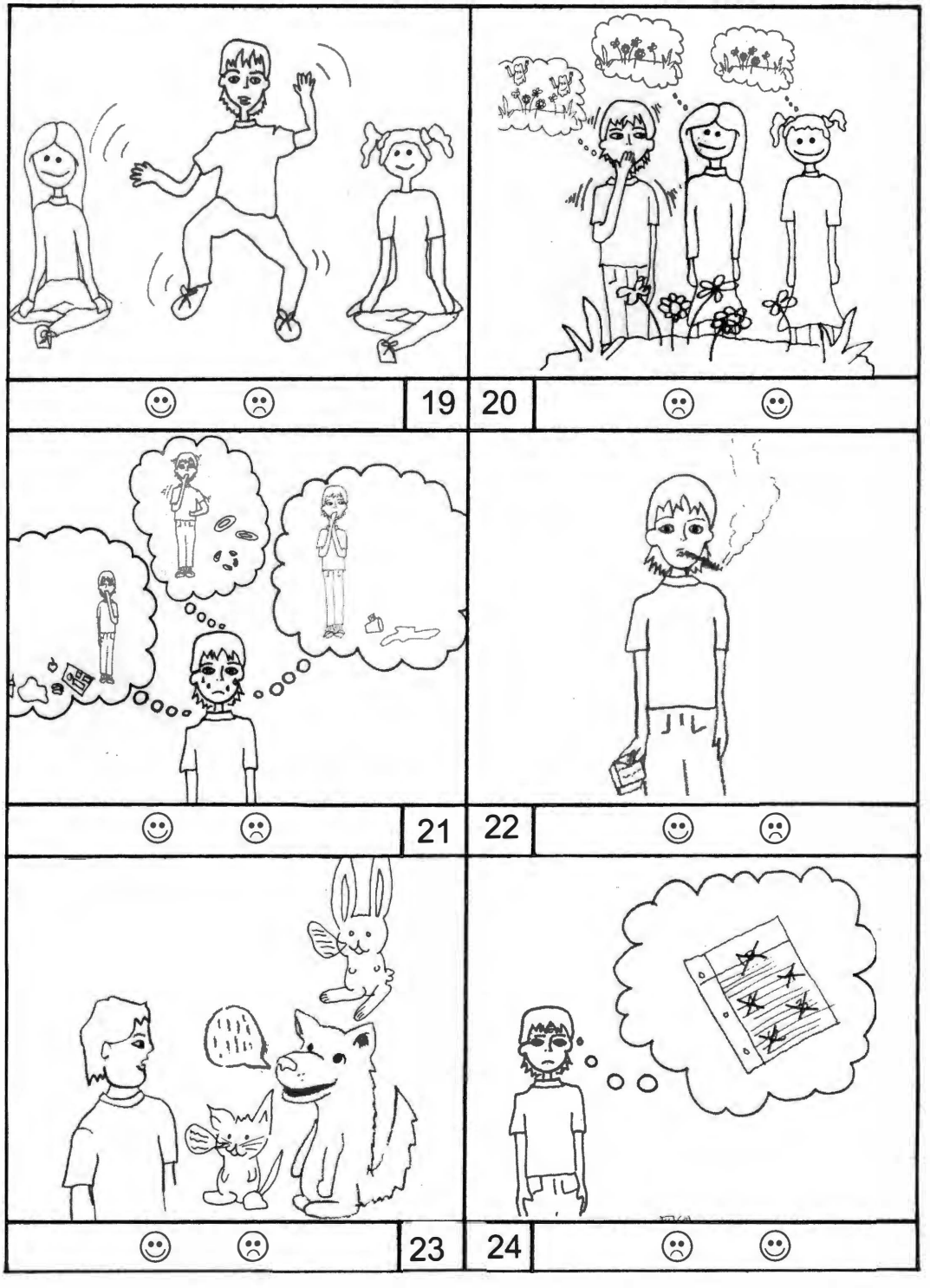
Appendix A. Continued.



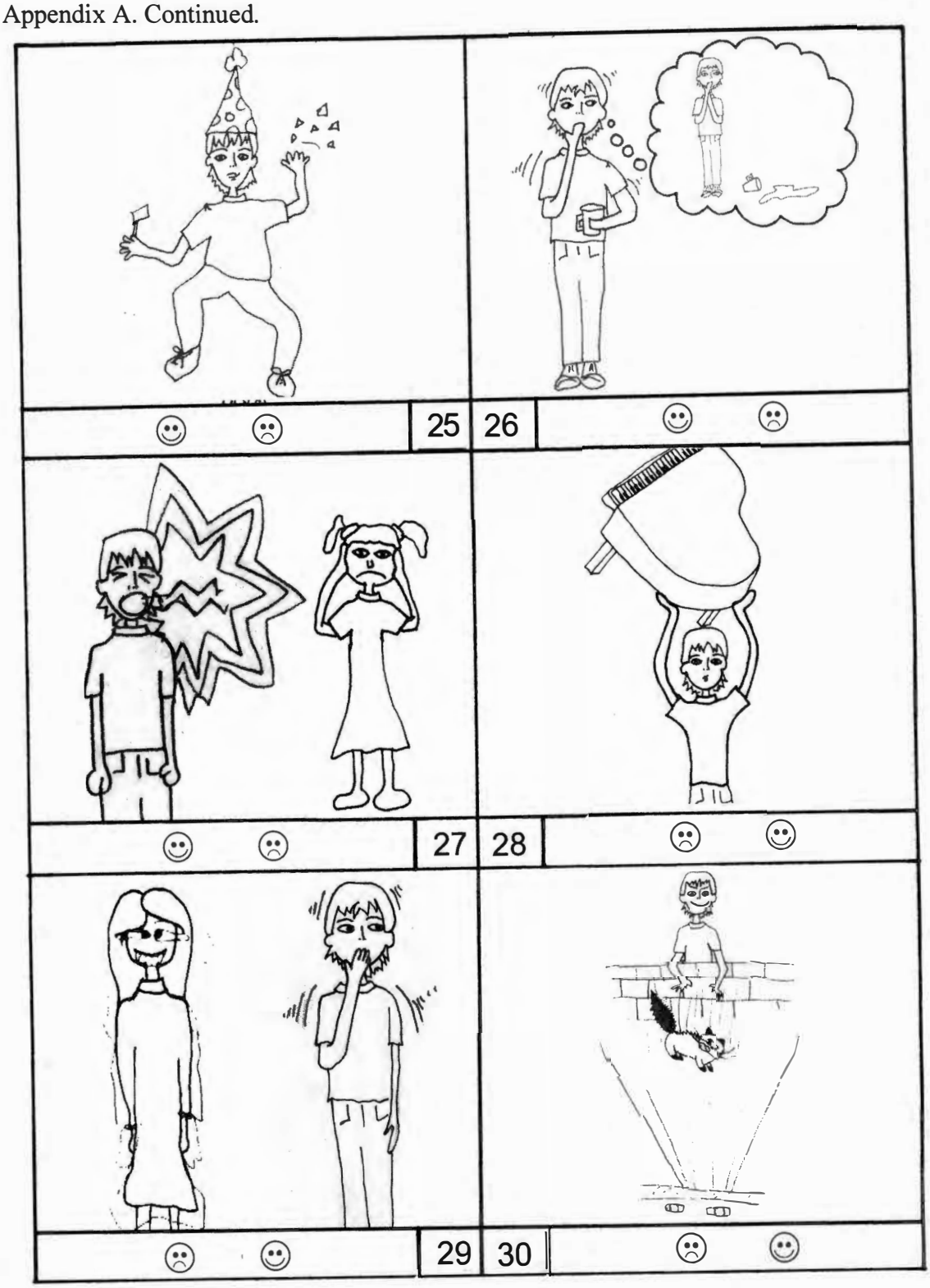
Appendix A. Continued.



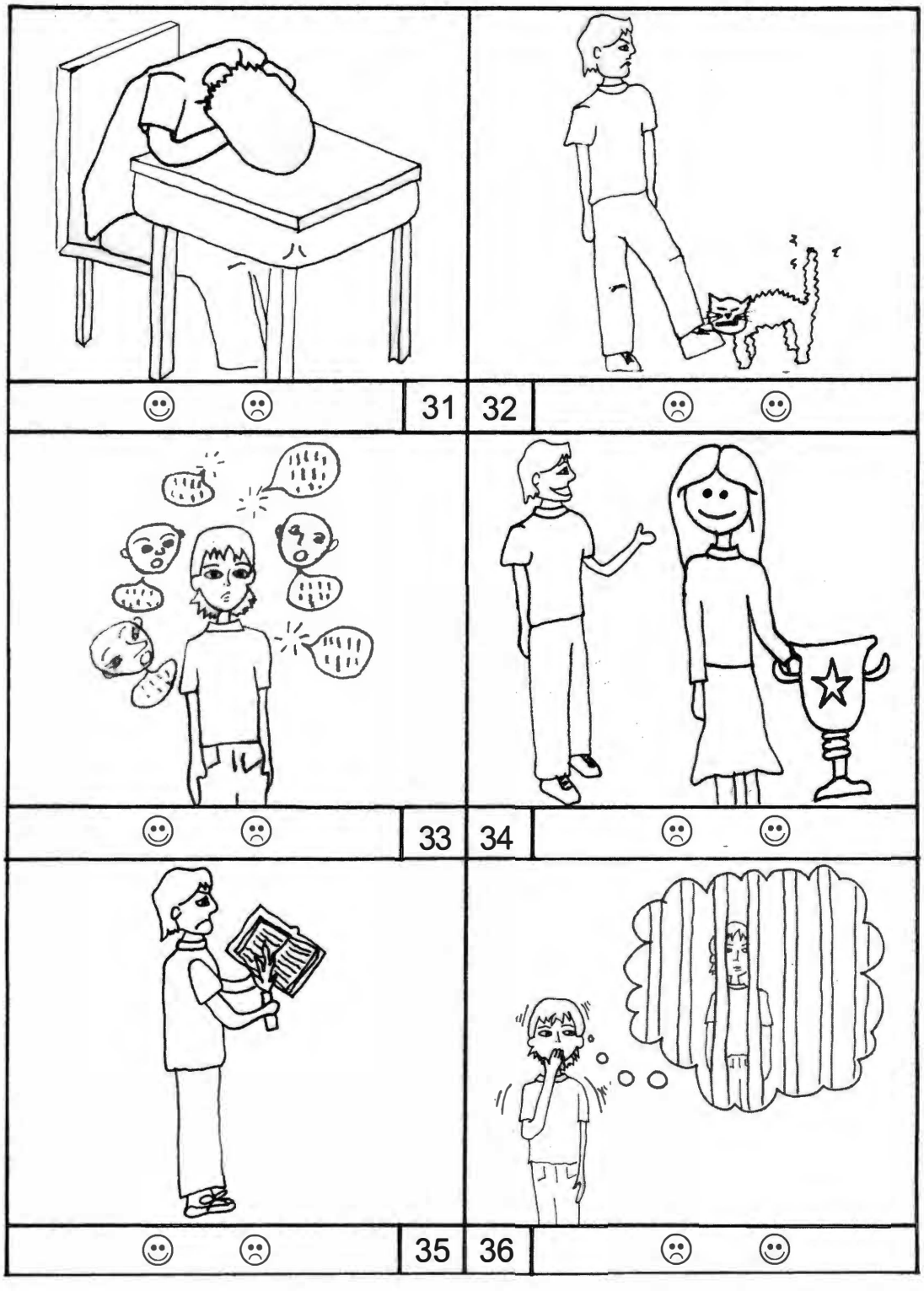
Appendix A. Continued.



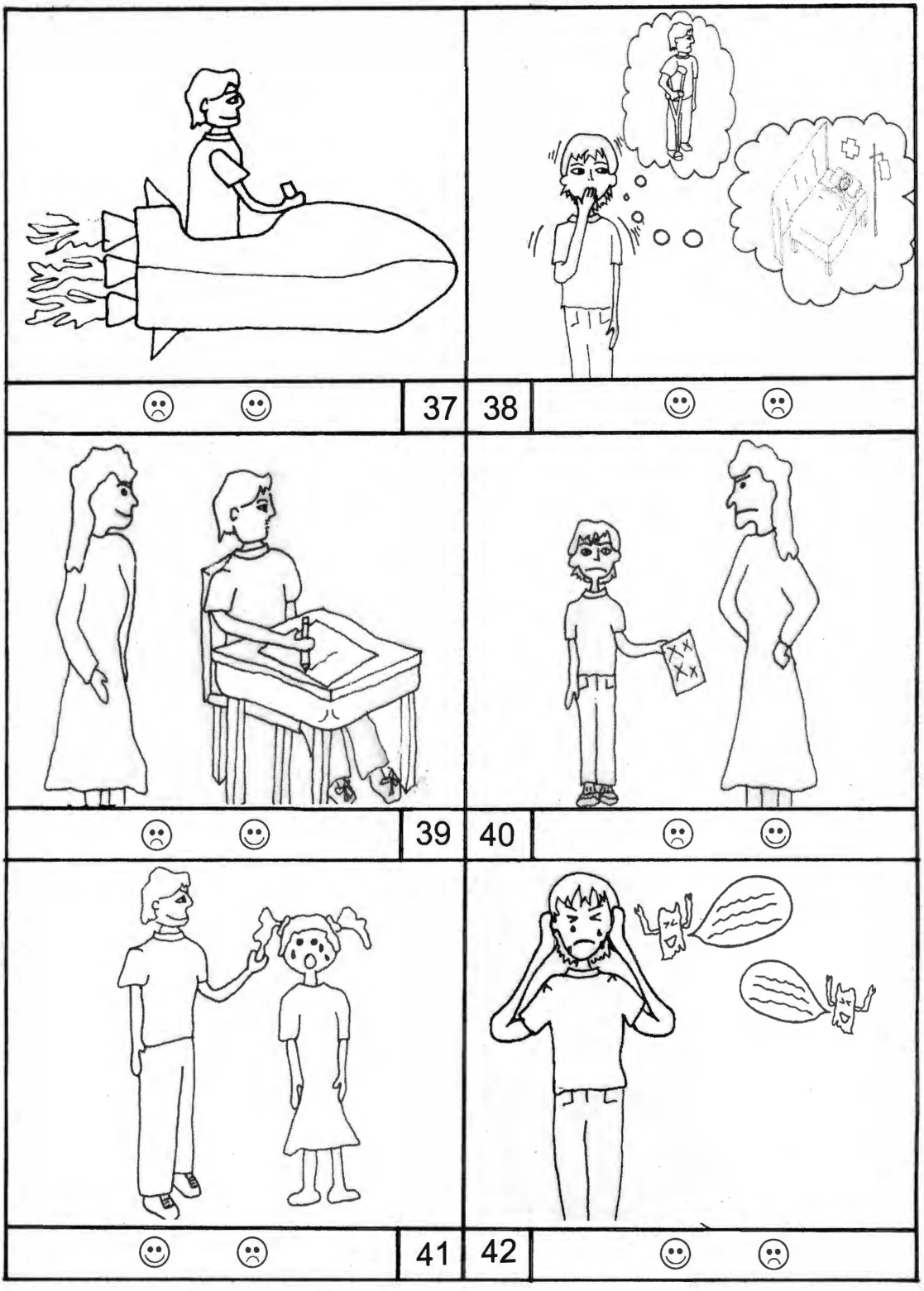
Appendix A. Continued.



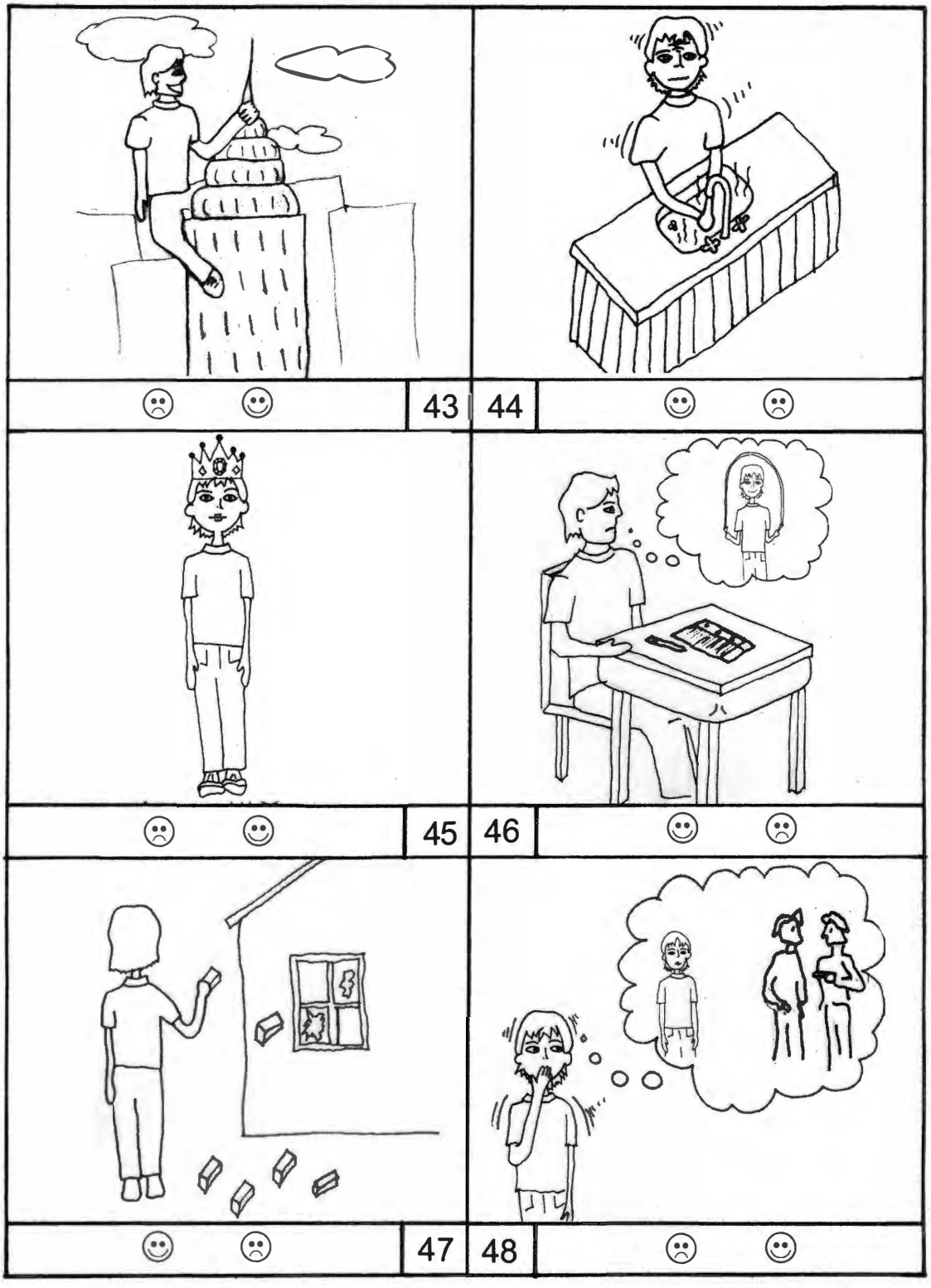
Appendix A. Continued.



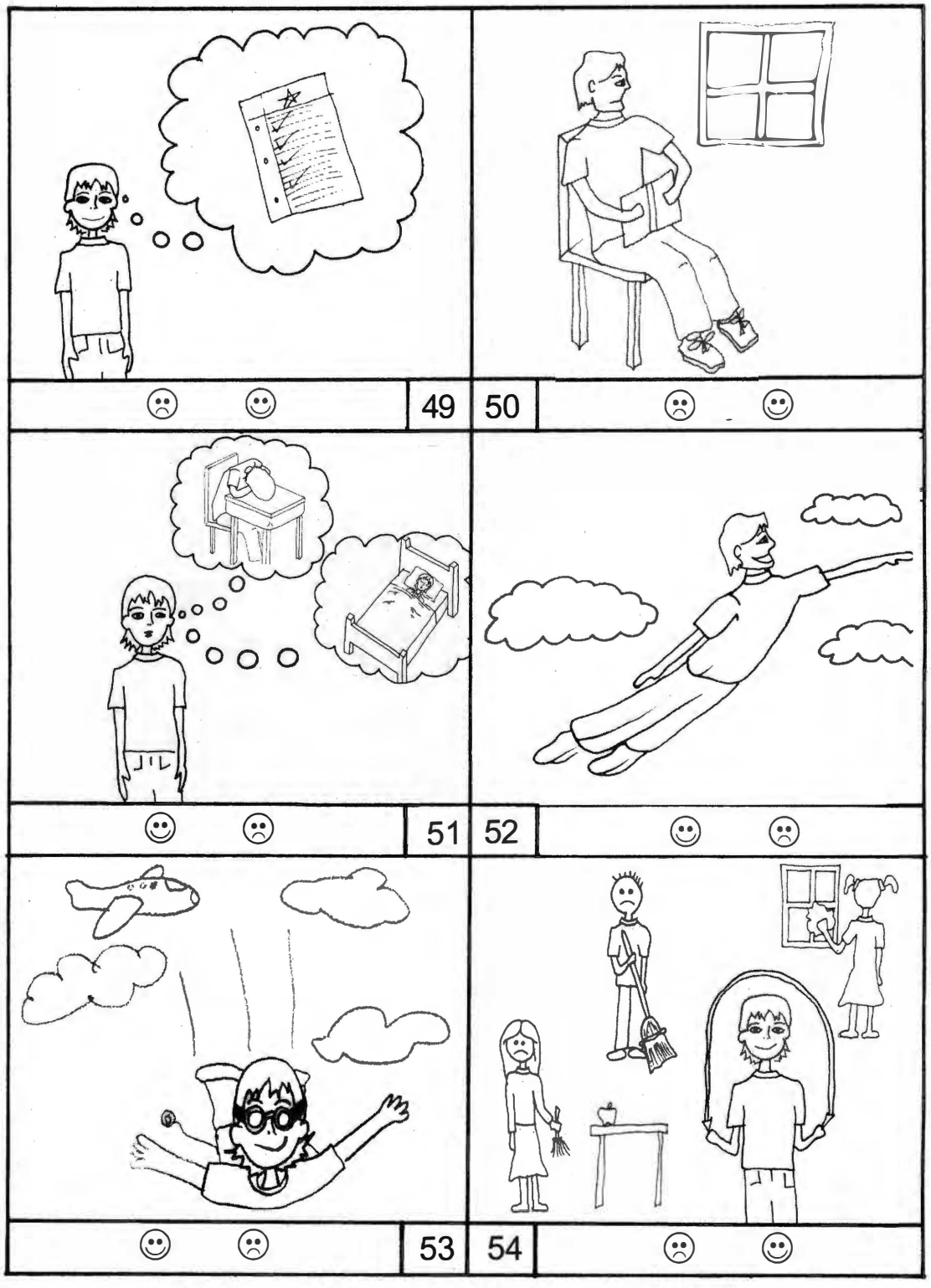
Appendix A. Continued.



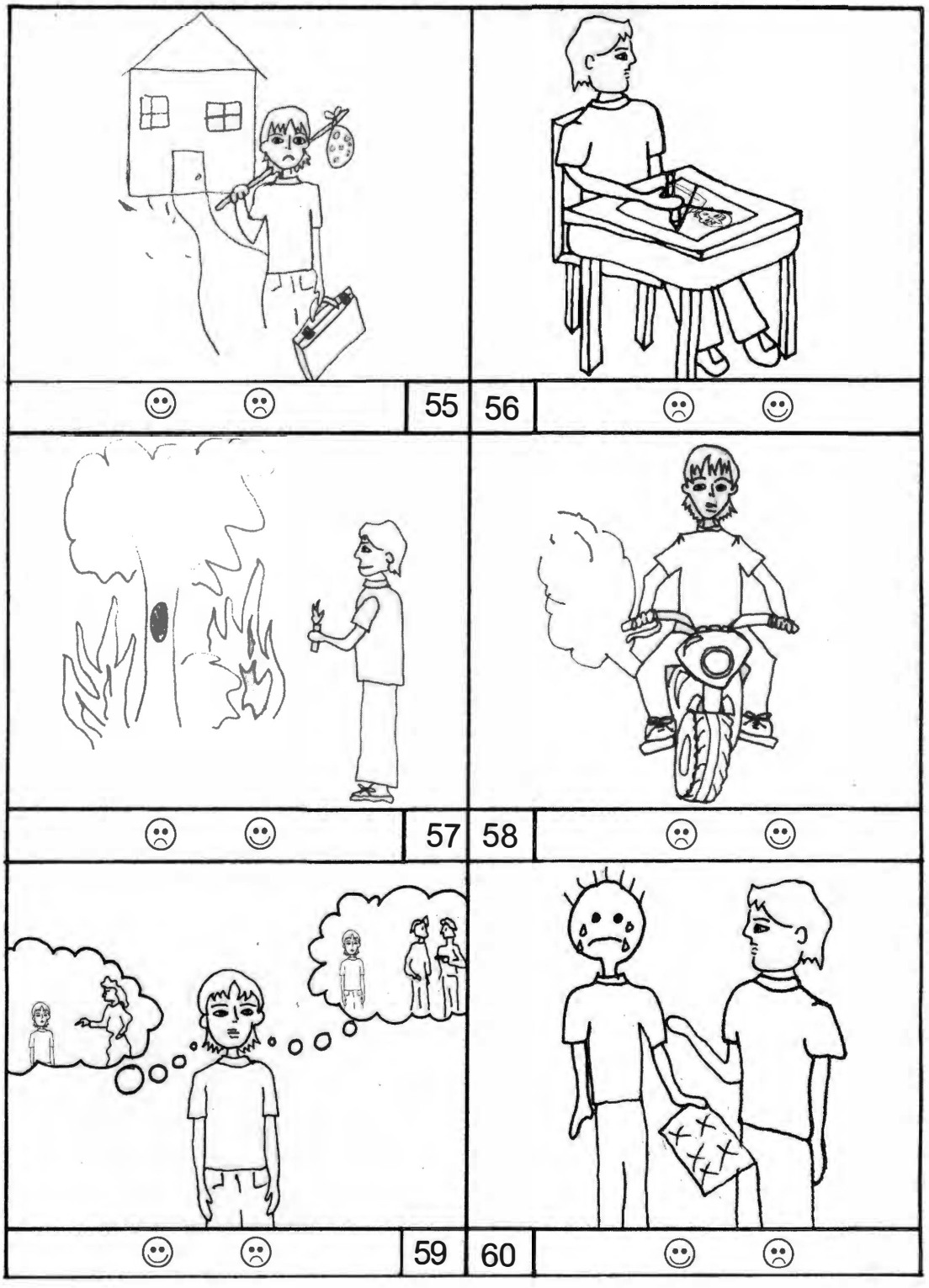
Appendix A. Continued.



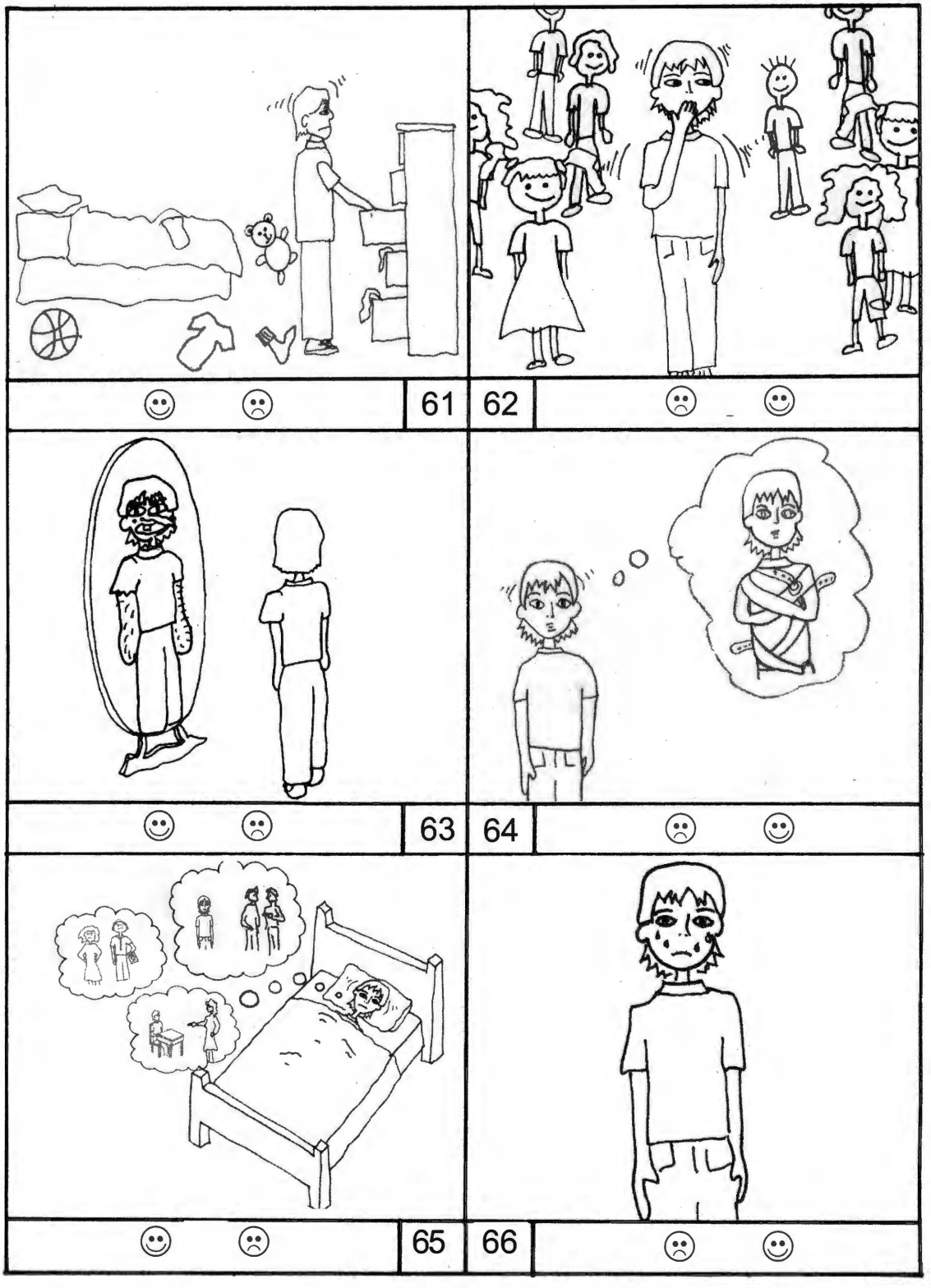
Appendix A. Continued.



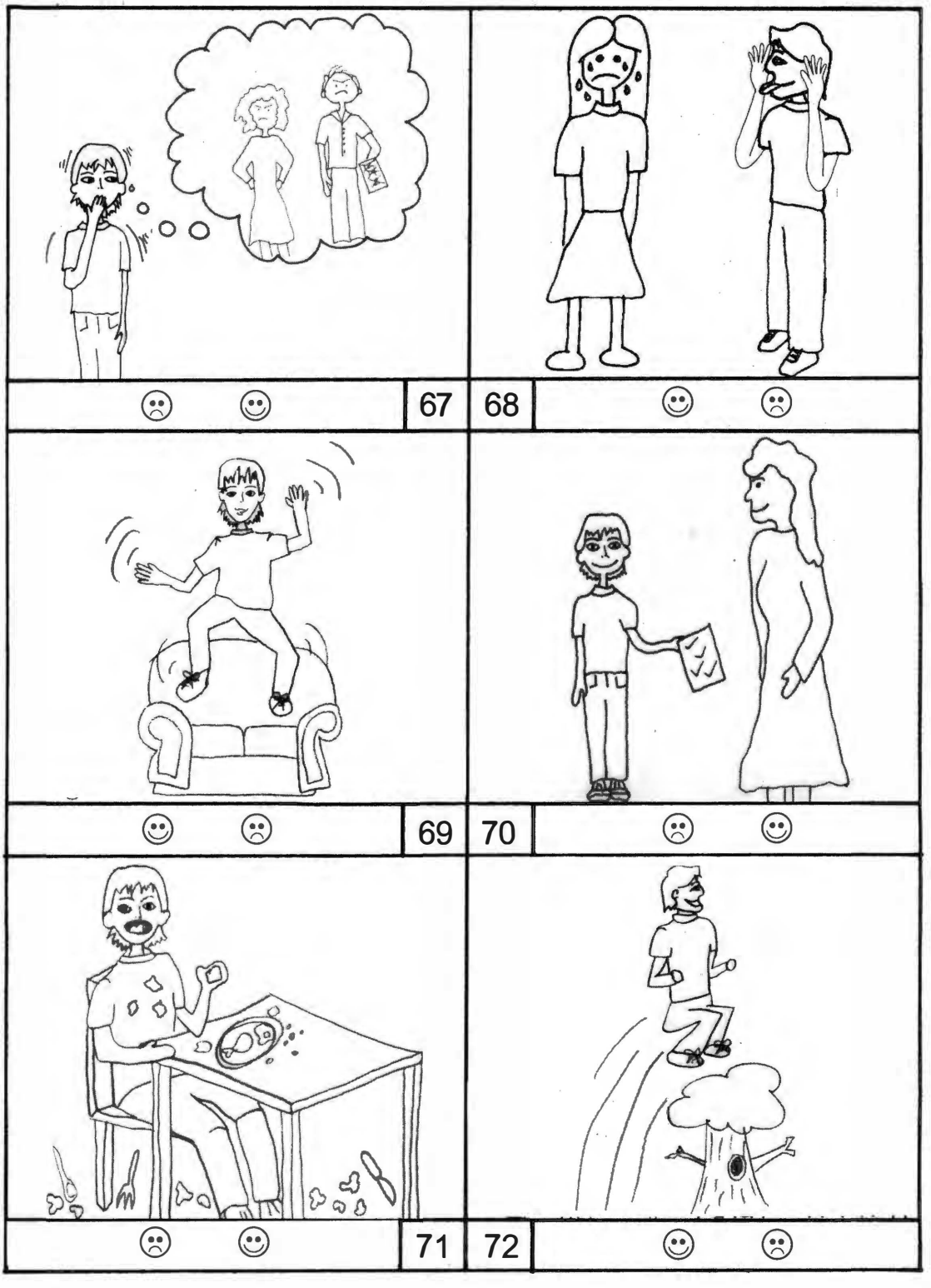
Appendix A. Continued.



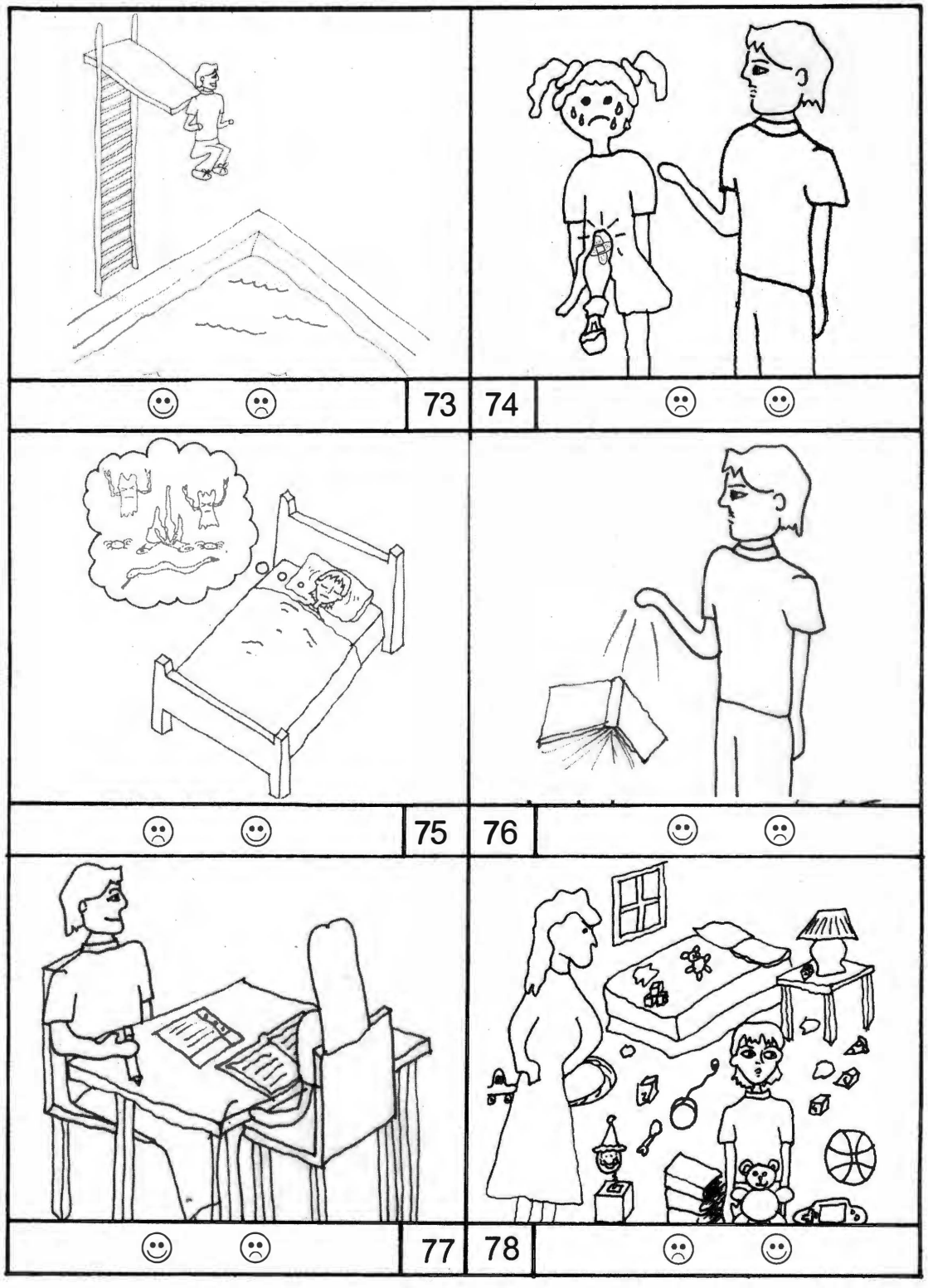
Appendix A. Continued.



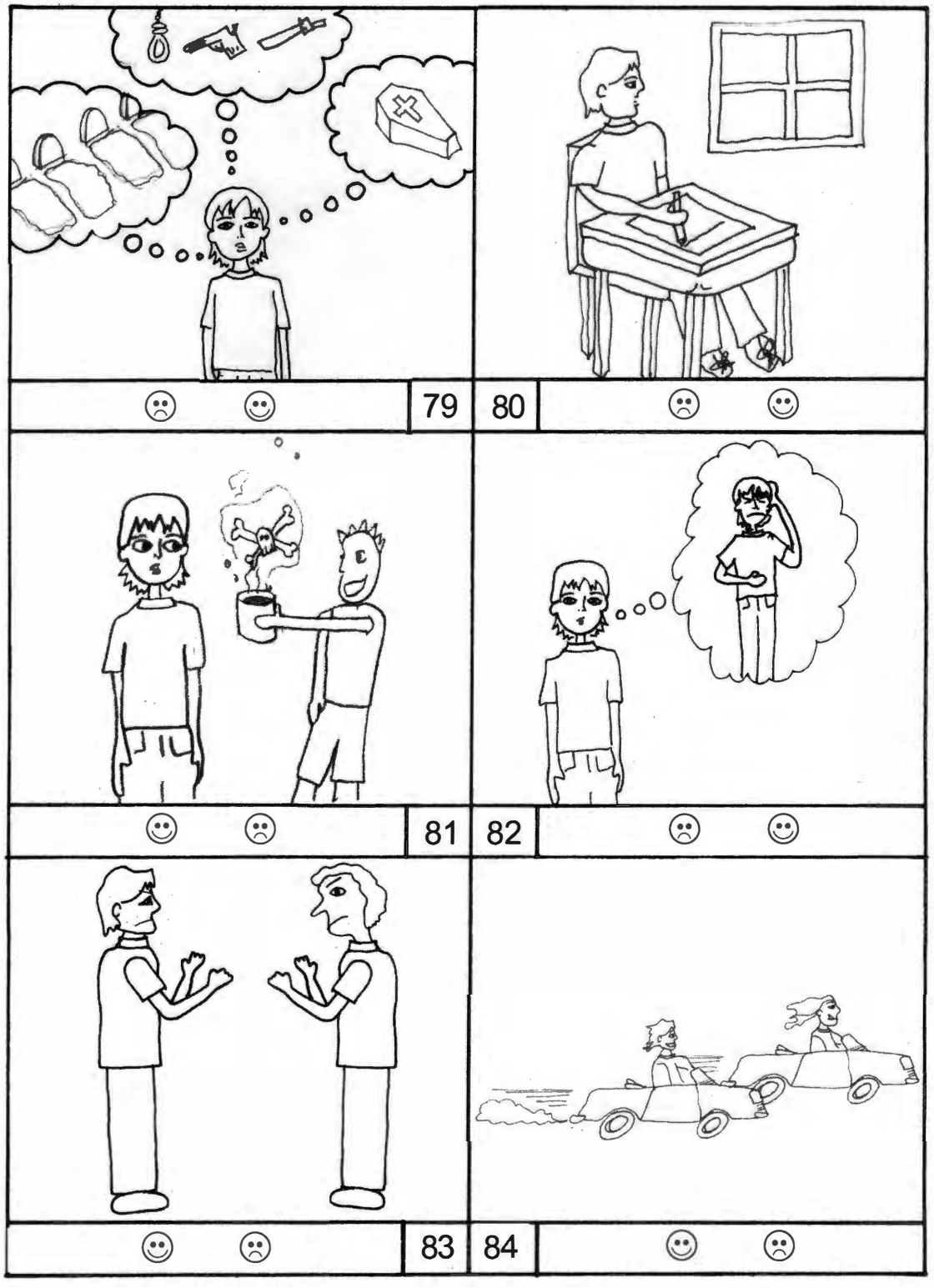
Appendix A. Continued.



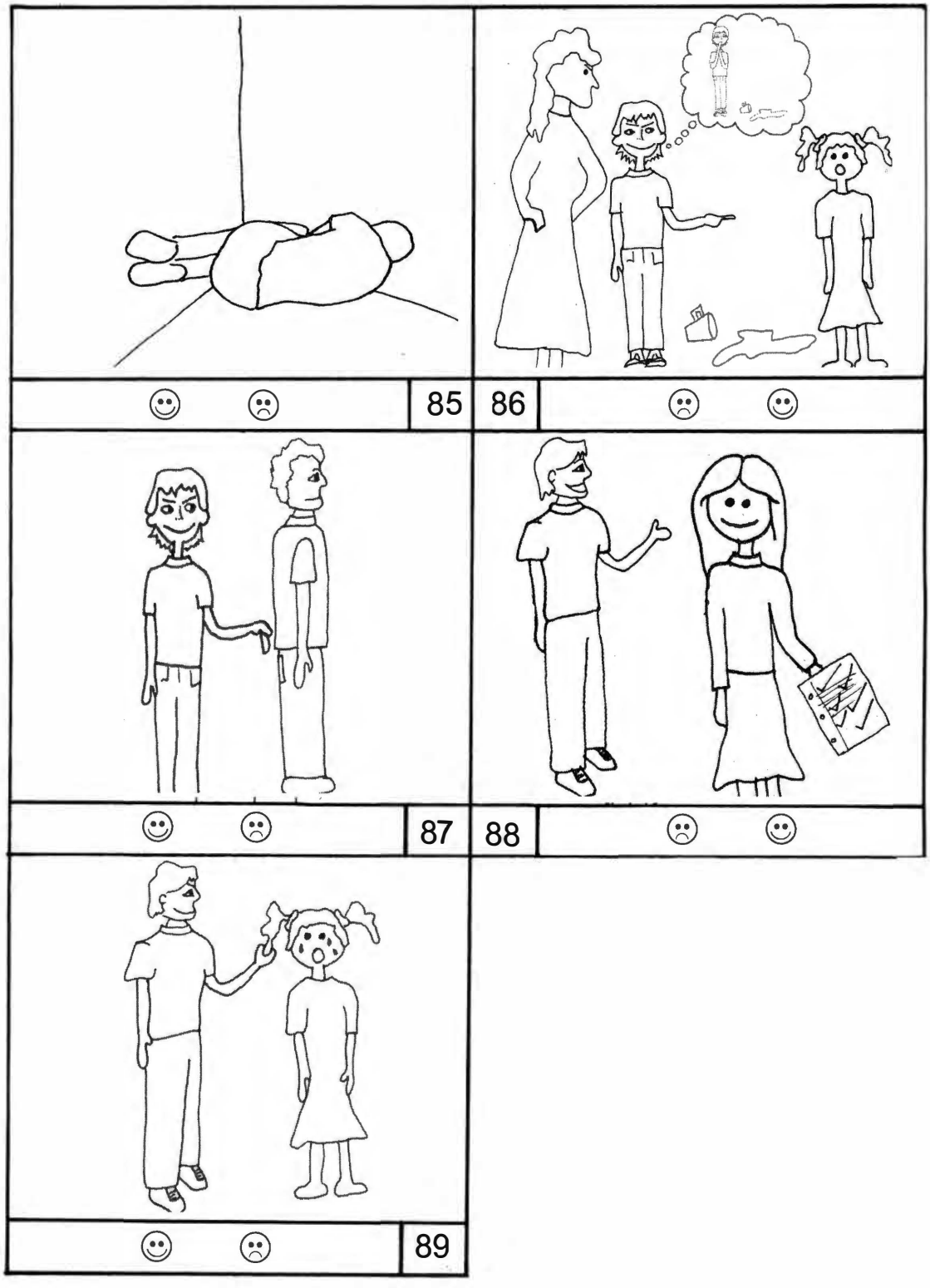
Appendix A. Continued.



Appendix A. Continued.



Appendix A. Continued.



Appendix B – UNIP Item Interpretation and Scale Assignment

Appendix B

UNIP Item Interpretation and Scale Assignment

Item Number	Item Interpretation	Scale(s)
1	I am a happy person.	Depression
		Consistency
2	It upsets me when I hear voices in my head.	Depression
		Consistency
3	I do not share toys with others.	Social Maladjustment
4	I get upset when others laugh at me.	Depression
5	I worry that I will do something bad.	Anxiety
6	I worry.	Anxiety
7	I eat messily.	ADHD
8	I do not pay attention to what I am doing.	ADHD
9	I worry about making bad grades.	Anxiety
		Academic Problems
		Consistency
10	I often am confused.	Depression
11	I am a happy person.	Depression
12	I am a king/queen.	Atypicality
		Consistency

Appendix B. Continued.

13	I have the ability to leap over trees.	Atypicality
		Consistency
14	I worry about getting in trouble with my teacher.	Anxiety
		Academic Problems
		ADHD
15	I daydream in class.	ADHD
16	I make good grades at school.	Academic Problems
		Consistency
17	I get in trouble when I make bad grades.	Anxiety
		Academic Problems
		Consistency
18	I am sad.	Depression
		Consistency
19	It is hard for me to sit still.	ADHD
20	I see things that others cannot see.	Atypicality
21	I worry about having accidents.	Anxiety
		Depression
22	I smoke.	Conduct Problems
		ADHD
23	I see and hear weird things	Atypicality

Appendix B. Continued.

24	I worry about making bad grades.	Anxiety
		Academic Problems
		Consistency
25	I am a happy person.	Depression
		Consistency
26	I worry when things do not go right for me.	Anxiety
27	I yell at others.	Conduct Problems
		Social Maladjustment
28	I can lift a piano over my head.	Atypicality
29	I worry about seeing weird things.	Anxiety
30	I like to hurt animals.	Conduct Problems
31	I put my head on my desk because I am sad.	Depression
32	I have fits of temper.	Conduct Problems
33	I get anxious when I hear voices in my head.	Anxiety
34	I congratulate others when they do good things.	Social Maladjustment
35	I destroy books.	Conduct Problems
36	I worry that bad things will happen to me.	Anxiety
37	I drive a rocket.	Atypicality
38	I worry about something bad happening to me.	Anxiety
39	I do not pay attention to my teacher.	ADHD

Appendix B. Continued.

40	I get in trouble when I make bad grades.	Anxiety
		Academic Problems
		Consistency
41	I do not get along with others.	Consistency
42	It upsets me when I hear voices in my head.	Depression
		Consistency
43	I can climb a skyscraper.	Atypicality
44	I do things over and over and cannot stop.	Anxiety
45	I am a king/queen.	Atypicality
		Consistency
46	I do not pay attention to my assignments.	ADHD
47	I do impulsive things.	ADHD
	I destroy things.	Conduct Problems
48	I worry about other people talking about me.	Anxiety
49	I make good grades at school.	Academic Problems
		Consistency
50	I am easily distracted.	ADHD
51	I wish I could sleep all the time.	Depression
52	I can fly.	Atypicality
53	I skydive.	Atypicality

Appendix B. Continued.

54	I do not help with classroom chores.	Social Maladjustment
55	I want to run away from home.	Anxiety
		Depression
56	I hate myself.	Depression
57	I set fire to trees.	Conduct Problems
58	I like to do exciting and dangerous things.	ADHD
59	No one likes me.	Anxiety
		Depression
60	I help people when the need assistance.	Social Maladjustment
61	I do not keep my room clean.	ADHD
62	I worry when I am in big crowds.	Anxiety
63	I am ugly. I hate my appearance.	Depression
64	I am scared of losing my mind.	Anxiety
65	I worry when I go to sleep at night.	Anxiety
66	I am sad.	Depression
		Consistency
67	I worry about getting in trouble with my parents.	Anxiety
		Depression
68	I tease and make fun of other people.	Social Maladjustment
69	I cannot sit still.	ADHD

Appendix B. Continued.

70	I make good grades.	Academic Problems
71	I have bad table manners.	Conduct Problems
		Social Maladjustment
72	I have the ability to leap over trees.	Atypicality
		Consistency
73	I jump off of really tall things.	ADHD
74	I comfort others when they are sad.	Social Maladjustment
75	I have nightmares.	Anxiety
76	I throw things when I am upset.	Social Maladjustment
77	I like to do my schoolwork with other people.	Social Maladjustment
78	I get in trouble when I do not clean my room.	Depression
		ADHD
79	I often think about death.	Depression
80	I cannot concentrate on my work.	ADHD
81	Someone has been trying to poison me.	Depression
82	I worry about aches and pains.	Anxiety
		Depression
83	I get in fights with other people.	Conduct Problems
		Social Maladjustment
84	I like to ride in cars that go fast.	ADHD

Appendix B. Continued.

85	Sometimes I am so sad that I want to be alone.	Depression
86	I do not admit when things are my fault.	Social Maladjustment
87	I steel things.	Conduct Problems
88	I congratulate others when they do well.	Social Maladjustment
89	I do not get along with others.	Consistency

Appendix C – Administration of the UNIP

Appendix C

Administration of the UNIP

Use of Gestures

Administration of the UNIP requires the use of eight rather universal gestures: head nodding up and down for "yes," head shaking from side to side for "no," pointing with the index finger (i.e., to the stimulus materials then to the examine) indicates "you do it now," palm rolling (i.e., a rolling motion with one hand, palm rotating quickly) to indicate "go ahead" or "hurry," open-handed shrug (i.e., lifting of the shoulders with palms up and a questioning facial expression) indicates "what is the answer" or "I don't know," hand waving (i.e., waving the hand horizontally over the stimulus material, palm up) indicates that the examinee is to attend to the material presented, stop (i.e., holding the open hand out in a nearly vertical position with the palm toward the examinee) indicates "stop," and finally, thumbs up (i.e., closing the fist and extending the thumb upward) indicates "good work," or acceptance, and conveys much the same message as a head nod. These gestures are employed by examiners who use the Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, 1998).

Use of "IP"

Each of the stimulus figures uses the same main character named IP. IP is an androgynous child about 9 or 11 years of age of no particular ethnicity. The examiner's job is to convey to the examinee that the examinee is represented by IP on each item and that the examinee is to indicate whether the behavior portrayed on each item is behavior the examinee typically exhibits (i.e., to what extent is IP "like me"). Examinees respond

by choosing from two response options: like me and not like me. The response options are represented by two illustrations of facial expressions (i.e., smile and frown) rather than words to avoid the complications associated with written response scales. Obviously it is very difficulty to convey examinee task demands nonverbally. Consequently, several demonstration and sample items are used until the examiner is certain the examinee understands what she/he is to do. In addition to the relatively universal gestures mentioned above, the examiner can use the large 8 X 10 demonstration and sample item stimulus plates and an 8 X 10 stimulus plate of IP for individualized administration. The same materials are used for group administration, except the stimulus plates are larger (11 X 17). Ultimately the examinee is to complete the circles in the Response Booklet showing one of the two "faces" in response to each of the 89 items.

Directions for administration of the UNIP

By using the directions below the examiner conveys to the examinee(s) that she/he/they is/are to indicate whether the behavior shown by IP is behavior that is typical for him/her/them. The directions are appropriate for group or individual administration, with minor modifications.

Demo Item 1.

1. The examiner shows the examinee(s) a picture of IP, presented on an 8 X 10 or 11 X 17 laminated stimulus plate. The examiner points to the picture of IP and then to herself/himself. Again, the examiner points to the picture of IP, then to himself/herself, then turns the picture of IP face down.

2. The examinee(s) is/are then shown the first demonstration item (a picture of IP studying). The examiner points to the picture of IP studying, then to himself/herself, then to the smiley face underneath the picture of IP studying, then uses the open-handed shrugging gesture for "I don't know" or "what is the answer" while engaging in eye contact with the examinee(s). After about 5 seconds the examiner nods her/his head "yes," and says, "like me."
3. After about 5 seconds, the examiner then points to the plate showing IP studying, to himself/herself, to the frowny face underneath the picture, and offers the "what is the answer gesture" for 5 seconds. The examiner then shakes her/his head "no," and says, "not like me."
4. After about 5 seconds, the examiner then points to the picture of IP studying, to himself/herself, then to each of the two options in turn, beginning with the smiley face, then the frowny face, nodding her/his head "yes" at the smiley face and says, "like me," then shaking her/his head "no" at the frowny face and says, "not like me."
5. The examiner then takes the red washable marker, points to the picture of IP studying, uses the open-handed shrug to indicate "what is the answer," and then very deliberately fills in the circle with the smiley face. The examiner then points to the smiley face and the picture of IP studying and nods his/her head "yes," and says, "Like me. I study."

Sample Item 1.

1. The examiner shows the examinee(s) the picture of IP. The examiner points to IP on the picture, then to the examinee(s). Again, the examiner points to IP on the picture, then to the examinee(s). The picture of IP is then turned face down.
2. The examinee(s) is then shown the first sample item (the 8 X 10 or 11 X 17 of IP studying). The examiner points to the picture of IP studying, then to the examinee(s), then to the smiley face underneath the picture of IP studying, then uses the open-handed shrugging gesture for "I don't know" or "what is the answer" while engaging in eye contact with the examinee(s). After about 5 seconds the examiner nods her/his head "yes," and says, "like you" followed by the open-handed shrugging gesture for "what is the answer."
3. After about 5 seconds, the examiner then points to the picture of IP studying, to the examinee(s), to the frowny face underneath the picture, and offers the "what is the answer gesture." After about 5 seconds the examiner nods her/his head "no," and says, "not like you" followed by the open-handed shrugging gesture for "what is the answer."
4. After about 5 seconds, the examiner then points to the picture of IP studying, to the examinee(s), then to each of the two options in turn, beginning with the smiley face, then the frowny face, nodding her/his head "yes" at the smiley face and says, "like you," then shaking her/his head "no" at the frowny face and says, "not like you" followed the open-handed shrugging gesture for "what is the answer?"
5. The examiner point to the examinee(s) and says, "you do it."

Demo Item 2.

1. The examiner shows the examinee(s) a picture of IP, presented on an 8 X 10 or 11 X 17 laminated stimulus plate. The examiner points to the picture of IP and then to herself/himself. Again, the examiner points to the picture of IP, then to himself/herself, then turns the picture of IP face down.
2. The examinee(s) is then shown the second demonstration item (a picture of IP riding a bicycle). The examiner points to the picture of IP riding a bicycle, then to himself/herself, then to the smiley face underneath the picture of IP studying, then uses the open-handed shrugging gesture for "I don't know" or "what is the answer" while engaging in eye contact with the examinee(s). After about 5 seconds the examiner nods her/his head "yes," and says, "like me."
3. After about 5 seconds, the examiner then points to the plate showing IP riding a bicycle, to himself/herself, to the frowny face underneath the picture, and offers the "what is the answer gesture" for 5 seconds. The examiner then shakes her/his head "no," and says, "not like me."
4. After about 5 seconds, the examiner then points to the picture of IP studying, to himself/herself, then to each of the two options in turn, beginning with the smiley face, then the frowny face, nodding her/his head "yes" at the smiley face and says, "like me," then shaking her/his head "no" at the frowny face and says, "not like me."
5. The examiner then takes the red washable marker, points to the picture of IP studying, uses the open-handed shrug to indicate "what is the answer," and then

very deliberately fills in the circle with the frowny face. The examiner then points to the frowny face and the picture of IP riding a bicycle and nods his/her head "no," and says, "Not like me. I do not ride a bicycle."

Sample Item 2.

1. The examiner shows the examinee(s) a picture of IP. The examiner points to IP on the picture, then to the examinee(s). Again, the examiner points to IP on the picture, then to the examinee(s). The picture of IP is then turned face down.
2. The examinee(s) is/are then shown the second sample item (the 8 X 10 or 11 X 17 of IP riding a bicycle). The examiner points to the picture of IP riding a bicycle, then to the examinee(s), then to the smiley face underneath the picture of IP studying, then uses the open-handed shrugging gesture for "I don't know" or "what is the answer" while engaging in eye contact with the examinee(s). After about 5 seconds the examiner nods her/his head "yes," and says, "like you" followed by the open-handed shrugging gesture for "what is the answer."
3. After about 5 seconds, the examiner then points to the picture of IP riding a bicycle, to the examinee(s), to the frowny face underneath the picture, and offers the "what is the answer gesture." After about 5 seconds the examiner nods her/his head "no," and says, "not like you" followed by the open-handed shrugging gesture for "what is the answer."
4. After about 5 seconds, the examiner then points to the picture of IP riding a bicycle, to the examinee(s), then to each of the two options in turn, beginning with the smiley face, then the frowny face, nodding her/his head "yes" at the

smiley face and says, "like you," then shaking her/his head "no" at the frowny face and says, "not like you" followed the open-handed shrugging gesture for "what is the answer?"

5. The examiner point to the examinee(s) and says, "you do it."

Appendix D – Tables

Table D1

Descriptive Statistics of the Universal Nonverbal Inventory of Personality

UNIP Scale	Mean	Standard Deviation
Anxiety	3.9	4.0
Depression	3.9	3.9
Atypicality	3.6	2.8
Academic Problems	.65	1.3
Attention Deficit Hyperactivity Disorder	6.4	4.0
Conduct Problems	.80	1.5
Social Maladjustment	1.7	1.8

Table D2

Descriptive Statistics of the Behavior Assessment System for Children-Second Edition

BASC-II Scales	Raw Score Mean	Raw Score Standard Deviation	Standard Score Mean*	Standard Score Standard Deviation*
Anxiety	8.7	3.2	46	6.3
Depression	2.8	2.2	45	5.2
Atypicality	1.4	1.8	45	3.5
Hyperactivity	5.0	1.7	48	3.3
Attention Problems	3.7	1.5	43	4.4
Interpersonal Relations	1.2	1.2	10	3.1
Attitude to School	2.9	1.5	40	3.9
Attitude to Teachers	2.9	1.8	41	4.2

* BASC-II Means and Standard Deviations can be compared to that of the general population (Mean = 50, Standard Deviation = 10).

Table D3

Internal Consistency Reliability Estimates (Cronbach's Alphas) of the Universal Nonverbal Inventory of Personality Scales

Scale	Cronbach's Alpha Reliability Coefficient
Anxiety	.85
Depression	.84
Atypicality	.78
Academic Problems	.77
Attention Deficit Hyperactivity Disorder	.75
Conduct Problems	.73
Social Maladjustment	.64

Table D4

Percentage of Inconsistent Responding to Universal Nonverbal Inventory of Personality Consistency Pairs

Total Inconsistencies	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	48	36.92	48	36.92
1	45	34.62	93	71.54
2	27	20.77	120	92.31
3	8	6.15	128	98.46
4	2	1.54	130	100.00

Table D5

Item-Total Correlations for the Universal Nonverbal Inventory of Personality Scales

Anxiety		Depression		Atypicality		Academic Problems	
Item No.	Item-Total Correlation	Item No.	Item-Total Correlation	Item No.	Item-Total Correlation	Item No.	Item-Total Correlation
5	.50	1	.25	12	.40	9	.51
6	.38	2	.45	13	.48	14	.31
9	.43	4	.54	20	.23	16	.51
14	.42	10	.30	23	.40	17	.46
17	.18	11	.23	28	.41	24	.61
21	.43	18	.57	37	.44	40	.42
24	.28	21	.29	43	.59	49	.57
26	.44	25	.35	45	.44	70	.65
29	.43	31	.36	52	.49		
33	.44	42	.48	53	.36		
36	.36	51	.31	72	.52		
38	.49	55	.36				
40	.33	56	.44				
44	.30	59	.33				
48	.42	63	.46				
55	.35	66	.55				
59	.38	67	.31				

Table D5 (continued)

Anxiety		Depression	
Item No.	Item-Total Correlation	Item No.	Item-Total Correlation
62	.44	78	.35
64	.53	79	.43
65	.30	81	.40
67	.41	82	.58
75	.52	85	.62
82	.66		

Table D5 (continued)

Attention Deficit Hyperactivity Disorder		Conduct Problems		Social Maladjustment	
Item No.	Item-Total Correlation	Item No.	Item-Total Correlation	Item No.	Item-Total Correlation
7	.51	22	.31	3	.21
8	.42	27	.33	27	.34
14	.34	30	.60	34	.28
15	.34	32	.48	54	.24
19	.32	35	.33	60	.30
22	.23	47	.29	68	.32
39	.23	57	.42	71	.26
46	.44	71	.43	74	.25
47	.26	83	.39	76	.37
50	.29	87	.51	77	.22
58	.21			83	.25
61	.38			86	.38
69	.50			88	.38
73	.29				
78	.39				
80	.36				
84	.31				

Table D6

Correlation Coefficients between Universal Nonverbal Inventory of Personality Scales and Behavior Assessment System for Children-Second Edition Scales

UNIP Scale	BASC-II Scale	Pearson Product Correlation Coefficient
Depression	Depression	.62*
Anxiety	Depression	.57*
Attention Deficit Hyperactivity Disorder	Attention Problems	.40*
Anxiety	Anxiety	.37*
Academic Problems	Attitude to Teachers	.36*
Attention Deficit Hyperactivity Disorder	Hyperactivity	.34*
Conduct Problems	Interpersonal Relations	.32*
Social Maladjustment	Interpersonal Relations	.27*
Academic Problems	Attitude to School	.17
Atypicality	Atypicality	.04

*$p < .01$.

VITA

Holly Michelle Hutchins will be receiving her Ph.D. in Education with a concentration in school psychology. She attended schools in Franklin County, Tennessee and graduated with honors from Franklin County High School in 1997. She graduated Summa Cum Laude and received a Bachelor of Arts Degree in Psychology from The University of Tennessee, Knoxville in the spring of 2001. In the fall of that year she returned to The University of Tennessee to pursue her doctoral degree. She has worked part-time as a Child Developmental Therapist for Douglas Cooperative Preschool throughout her doctoral program. She completed an internship with Cherokee Health Systems in Morristown, Tennessee, where she served as a School Psychology Intern for the Hamblen County School System, a Mental Health Consultant for the Clinch-Powell Head Start Program, a Behavioral Health Consultant for a pediatric clinic, and assisted in autism evaluations. Holly's primary area of interest is in early intervention services for preschoolers.

4535 0545 24
11/09/05 MRB