7-28-2011

Detection of Dissimulation in Children on the BASC-2 Self-Report of Personality

Nicole A. Begg

To the University Council:

The Dissertation Committee for Nicole A. Begg certifies that this is the final approved version of the following electronic dissertation: "Detection of Dissimulation in Children on the BASC-2 Self-Report of Personality."

_____
Randy G. Floyd, Ph.D.
Major Professor

We have read this dissertation and recommend
its acceptance:

_____
Thomas K. Fagan, Ph.D.

_____
Elizabeth Meisinger, Ph.D.

_____
Meghan E. McDevitt-Murphy, Ph.D.

Accepted for the Graduate Council:

_____
Karen D. Weddle-West, Ph.D.
Vice Provost for Graduate Programs

DETECTION OF DISSIMULATION IN CHILDREN
ON THE BASC-2 SELF-REPORT OF PERSONALITY

by

Nicole A. Begg

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Psychology

The University of Memphis

August 2011

Abstract

Begg, Nicole A. Ph.D. The University of Memphis. August 2011. Detection of Dissimulation in Children on the BASC-2 Self-Report of Personality. Major Professor: Randy G. Floyd

Children are capable of deliberately distorting or misrepresenting psychological symptoms on self-reports of behavior and personality, which is referred to as dissimulation. Malingering and defensiveness are two forms of dissimulation that involve exaggerating symptoms or denying symptoms, respectively. The purpose of this study was to examine the extent to which the Behavior Assessment System for Children-2 Self-Report of Personality validity scales identified dissimulation response styles in children. This investigation employed a between-subjects experimental simulation design with three conditions. One hundred and eighteen children were randomly assigned to a control group or to one of two simulation groups: (a) a malingering group to simulate behavioral and emotional symptoms, or (b) a defensive group to simulate minimization or denial of behavioral and emotional symptoms. Specific research questions addressed whether the validity scales were useful in signaling caution for children who were instructed to malinger and for children who were instructed to respond defensively. Findings suggested that the F Index is most sensitive to detecting malingering. Results also indicated that the L Index may be a useful indicator of malingering, but this scale did not perform as well as expected in detecting children who were instructed to respond defensively.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1: Introduction

Detection of Dissimulation in Children

on the BASC-2 Self-Report of Personality

Self-reports of behavior and personality are valuable components in the

assessment of children. The usefulness of the knowledge gleaned from self-reports

depends upon their accuracy. Any assessment that includes children's perceptions of their

own behavior and personality characteristics may call into question the veracity of their

self-disclosure. Inconsistencies or questionable results are often attributed to children's

capacity to understand the directions. However, children are capable of deception (Lewis,

1993) and may minimize or exaggerate symptoms for a variety of reasons (Rogers,

1997a). Therefore, it is important that self-report instruments can reliably detect

dissimulation response patterns that potentially invalidate results.

Research has demonstrated that children have the capacity to engage in deceptive

strategies by the age of 2 years 6 months (Chandler, Fritz, & Hala, 1989) although their

ability to successfully deceive is limited based on their poor conceptual understanding of

false beliefs (Bussey, 1992; Wimmer, Gruber, & Perner, 1984). Lies told by young

children tend to be categorical in nature (i.e., yes/no; true/false), whereas lies told by

older children and adolescents tend to be more organized and intricate in detail (Bussey,

1992). As children develop in their cognitive skills, they become better equipped at

deception (Lewis, 1993). The prevalence rates of lying are found to be higher in children

referred for clinical treatment than in nonclinical populations (Behar, 1977; Ferguson,

Partyka, & Lester, 1974). Prevalence rates of lying stabilize or increase with age

(Achenbach & Edelbrock, 1981; Stouthamer-Loeber, 1986; Stouthamer-Loeber & Loeber, 1986).

Deception by "deliberately distorting or misrepresenting psychological symptoms" is referred to as *dissimulation* (Rogers, 1997b, p. 12), and it may take several forms including *malingering* and *defensiveness*. Malingering (American Psychiatric Association, 1994) or "faking bad" is a response style that is characterized by intentional fabrication or gross exaggeration of psychological or physical symptoms to give the appearance of being sick or poorly adjusted in order to realize secondary gain. Common incentives that might motivate a child to feign illness might include to seek attention or to avoid punishment or unpleasant tasks. According to the DSM-IV, malingering should be strongly suspected in cases of forensic referral for evaluation, discrepancy between reported symptoms and assessment results, failure to comply or cooperate with evaluation or treatment, or diagnosis of antisocial personality disorder (American Psychiatric Association, 1994). Conversely, responding in a manner to intentionally minimize or deny symptoms is referred to as defensiveness or "faking good." Defensiveness (Rogers, 1984), as a response style, is directly opposite to malingering and is also motivated by external incentives. Children may attempt to present themselves in the most favorable light in order to avoid getting into trouble or to seek praise or adulation for being a "good girl" or "good boy."

Dissimulation Research with Adolescents

Despite children's capacity for and prevalence rates of lying and deception, there is a paucity of research on dissimulation on self-report instruments in this age group. Prior to the last decade, information about malingering and defensiveness has been

extrapolated from adult studies of dissimulation or the deception literature, which are inadequate for understanding deceptive processes in younger populations (McCann, 1998; Rogers, 1997a). Although published investigations of preadolescents and younger children appear to be absent, there are a number of studies that have been conducted with adolescents.

The simulation design has been the methodology of choice in the study of dissimulation in adolescents. A simulation study (Rogers, 1997c) is an analogue design in which those participants randomly assigned to an experimental condition are instructed to adopt a desired response style (e.g., malingering, defensive responding) during completion of an assessment tool. Some investigators offer incentives to those participants that simulate most effectively. Participants who are randomly assigned to the control condition are asked to respond in a truthful manner and are provided with standard instructions for the assessment tool. A distinct advantage of the simulation design is that it permits well-controlled manipulation of experimental conditions that neutralize threats to internal validity.

The following review provides a summary of relevant work involving adolescents as shown in Table 1. Of 11 studies reviewed, eight were conducted using the Minnesota Multiphasic Personality Inventory-Adolescent (MMPI-A; Butcher et al., 1992) or its precursor, the MMPI (Hathaway & McKinley, 1940). Prior to the development of the MMPI-A in 1992, the MMPI was widely used for assessing adolescents (Archer, Maruish, Imhof, & Piotrowski, 1991) despite the lack of an age-appropriate normative sample. An overview of these scales and their derivations is provided followed by the literature review of malingering and defensiveness research with adolescents.

Table 1

Summary of Dissimulation Research with Adolescents

Authors	Participants	Materials	Study Focus
Archer, Gordon, & Kirchner (1987)	94 nonclinical & 24 clinical	MMPI	Malingering and Defensiveness
Herkov, Archer, & Gordon (1991)	283 nonclinical & 119 clinical	MMPI	Malingering and Defensiveness
Stein, Graham, & Williams (1995)	138 nonclinical & 138 clinical	MMPI-A	Malingering
Rogers, Hinds, & Sewell (1996)	53 offenders receiving clinical treatment	MMPI-A SIRS SIMS	Malingering
Lucio, Durán, Graham, & Ben-Porath (2002)	146 nonclinical Mexican & 146 clinical Mexican	MMPI-A *Spanish Version*	Malingering
Baer, Ballenger, & Kroll (1998)	26 nonclinical & archived nonclinical cases- *number not reported* & 48 clinical	MMPI-A	Defensiveness
Stein & Graham (1999)	137 nonoffenders & 140 offenders	MMPI-A	Defensiveness
Wrobel et al. (1999)			
Study #1	30 clinical	PIY	Defensiveness
Study #2	232 clinical	PIY, PPVT-R, Experimental adjective checklist	Not applicable
Study #3	108 nonclinical	PIY	Malingering

(table continues)

Table 1 (Continued)

Summary of Dissimulation Research with Adolescents

Authors	Participants	Materials	Study Focus
Stein & Graham (2005)	126 substance and non-substance abusing offenders	MMPI-A	Defensiveness

The original MMPI played a prominent role in the history of personality assessment and consisted of 504 test items designed to provide psychiatric diagnostic information (Anastasi & Urbina, 1997), and it was later revised, restandardized, and published as the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). The MMPI-A is the 478-item version of the MMPI-2 designed for use with 14 to 18 year olds (Graham, 2000; Sattler, 2002). Three validity scales on the MMPI-A, Lie (L), Infrequency (F), and Correction (K), were maintained from the original MMPI, but changes were made to the F scale. Additional validity scales that were developed for the MMPI-2 and included on the MMPI-A are the Variable Response Inconsistency (VRIN) and True Response Inconsistency (TRIN) scales. The L scale was developed to assess blatant attempts to deny minor shortcomings (e.g., "I do not read the instruction manual in its entirety prior to attempting to use new appliances") that most individuals would be likely to admit. The F scale includes items rarely endorsed by most individuals (e.g., "I have never seen a television set") and was designed to assess highly unusual response patterns. The F scale was broken down into two additional scales (Infrequency 1 [F1] and Infrequency 2 [F2]), which provide information about the first and second parts of the instrument, respectively. The K scale was designed for the same purpose as the F scale but items are far more difficult for sophisticated test-takers to identify due to subtle

content. The VRIN scale was designed to detect inconsistency in responding on pairs of items with similar or opposite content. The TRIN scale was designed to assess indiscriminant responding of true or false on pairs of items with opposite content. The MMPI-A includes 10 standard clinical scales retained from the original MMPI (i.e., Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Masculinity-Femininity, Paranoia, Psychasthenia, Schizophrenia, Hypomania, and Social Introversion) and 15 new content scales (i.e., Anxiety, Obsessiveness, Depression, Health Concerns, Alienation, Bizarre Mentation, Anger, Cynicism, Conduct Problems, Low Self-Esteem, Low Aspiration, Social Discomfort, Family Problems, School Problems, and Negative Treatment Indicators).

Research on malingering. Early work conducted by Archer et al. (1987) found that when researchers instructed adolescents to malinger serious emotional or psychological problems on the MMPI, malingering was easily detected on resulting profiles. Symptoms were grossly exaggerated and there were marked elevations on the F scale and elevations within the clinical range on the clinical scales. Significant effects were found for race and gender in the malingering condition. Results indicated that White adolescents produced more (a) marked elevations on 7 of 10 clinical scales and the F scale and (b) lower scores on the L and K scales than Black adolescents, thus yielding a more exaggerated representation of symptomatology. Also, females produced more exaggerated symptom profiles with higher elevations on four clinical scales and the F scale; and lower scores on the L scale.

Herkov, Archer, and Gordon (1991) were specifically interested in how well the Wiener–Harmon subtle-obvious subscales could detect malingering on the MMPI.

6

Wiener (1948) created the subscales by differentiating items on six of the clinical scales that were either easy or difficult to identify as indicators of emotional or psychological distress. When comparing adolescent inpatients who received standard instructions and adolescent nonpatients who received instructions to simulate psychopathology, the best indicators of malingering were the F and K scales and the best indicator of defensiveness was the L scale. Results suggested that the subtle–obvious scales provided no incremental validity beyond the standard validity scales. It is noted that the subtle–obvious scales were deleted from later revisions of the MMPI based on several studies of adults that yielded similar findings (Graham, 2000).

After the development of the MMPI-A, which included a number of item changes and an age-appropriate normative sample, Stein, Graham, and Williams (1995) examined the utility of the MMPI-A validity scales in differentiating between nonclinical adolescents instructed to malinger and adolescents from both clinical and nonclinical settings instructed to respond according to standard protocol for the measure. Results were consistent with a previous study of adolescents using the original MMPI (Archer, Gordon, & Kirchner, 1987) in identifying malingerers and obtaining similar mean profiles. Findings yielded accurate differentiation by the F, F minus K, F1, and F2 scales between clinical and nonclinical adolescents who were instructed to malinger and between nonclinical adolescents instructed to malinger and nonclinical adolescents who were given standard instructions. He suggested that the difference between the F and K raw scores might yield a useful index of faking bad. Malingering yielded marked elevations on the F scale and yielded scores at or below 50 on the L and K scales. The raw F score was determined to be just as effective or better at detecting malingering as

other scales and indexes. Regarding the F minus K index, Gough (1950) noted that malingerers tended to obtain significantly higher F than K scale scores and that the difference in raw scores might be a useful indicator of malingering.

In an investigation of malingering psychopathology among adolescent offenders, Rogers, Hinds, and Sewell (1996) studied the clinical utility of the Structured Interview of Reported Symptoms (SIRS; Rogers, 1992; Rogers, Bagby, & Dickens, 1992), MMPI-A, and Screening Index of Malingered Symptoms (SIMS; Smith, 1992) in the assessment of faking bad. Findings suggested that the F-K > 20 as a screen for malingering could be useful but found that the frequently used F, F1, and F2 scales were ineffective indicators of malingering. Both the SIRS and the SIMS were found to be effective in distinguishing malingered protocols. Study findings suggested that including both the SIRS and the MMPI-A in an evaluation yielded incremental validity and the best chance to detect malingering in adolescents.

A Spanish version of the MMPI-A was translated and adapted for use with Mexican adolescents by Lucio (1998). Lucio, Durán, Graham, and Ben-Porath (2002) compared the profiles on nonclinical Mexican adolescents who were instructed to malinger and both clinical and nonclinical Mexican adolescents who were given standard instructions. They found that the F, F1, and F2 scales and the F-K index differentiated participants among all three groups, although higher cutoff scores were required to differentiate between girls than to differentiate between boys.

Whereas most research on malingering in adolescents focused on the MMPI and MMPI-A, Wrobel et al. (1999) investigated malingering using the Personality Inventory for Youth validity scales. High school students were asked to fake-moderate, fake-severe,

respond randomly, or respond accurately in completing the PIY. Higher mean scores were obtained on the PIY Validity scale (designed to assess excessively negative responding with the inclusion of items improbably or infrequently endorsed by most adolescents) and PIY Fake Bad scale (designed to detect deliberate attempts to appear emotionally disturbed) by participants in the fake-moderate, fake-severe, and random response groups than those in the accurate response group. Further, lower mean scores were obtained on the PIY Defensiveness scale (designed to detect attempts to present oneself in an unrealistically favorable light) by participants in both fake bad groups (moderate and severe) in comparison to participants in the random response group and accurate response group.

In summary, research suggests that adolescents can successfully fake bad when instructed to malinger. There is empirical support for the ability of several validity scales on the MMPI and MMPI-A to detect malingering. For example, elevated F scale scores from the MMPI forms consistently appear to be the best indicators of malingering, whereas lower K scores from the MMPI forms also appear to be good indicators of malingering.

Research on defensiveness. Other studies have investigated how well the validity scales from the MMPI, MMPI-A, and PIY can detect defensiveness when adolescents are instructed to minimize or deny symptoms of emotional or psychological distress. Archer et al. (1987) found that two kinds of profiles emerged when they instructed adolescent inpatients with a variety of psychiatric diagnoses to respond defensively on the MMPI. Individuals who were ineffective at simulating healthy adjustment were younger and had higher baseline scores on clinical scales upon admission, whereas those who were

9

effective at simulating healthy adjustment were older and had lower baseline scores on clinical scales upon admission. Thus, findings suggested that older adolescents who experience less psychiatric disturbance may be able to escape detection by producing MMPI-A profiles within normal limits, including the validity scales.

Herkov et al. (1991) compared adolescent patients who received instructions to respond defensively and adolescent nonpatients who received standard instructions. They found that they could accurately detect 71% of participants who were defensive and 90% of participants who received standard instructions by using a cutoff score of 55 or more on the L scale. Further, they found that the Wiener–Harmon subtle–obvious scales were ineffective in discriminating the defensive responders from the standard instruction responders.

In a study of defensiveness on the MMPI and MMPI-A utilizing clinical and nonclinical adolescent samples, Baer, Ballenger, and Kroll (1998) found that it was possible to detect defensiveness on adolescent profiles using the MMPI and MMPI-A using the L and K scales. Furthermore, findings indicated that the L scale was most effective in discriminating profiles of those instructed to respond defensively versus standard instructions.

Stein and Graham (1999) also conducted a study of defensiveness and found that the MMPI-A validity scales could distinguish between adolescent offenders who were instructed to respond defensively and both offenders and non-offenders who were instructed to respond to the standardized instructions. Slightly different cutoff scores were needed for the differentiations. Further, mean profiles for the adolescent offenders

10

were consistent with previous findings for adults on the MMPI and MMPI-2 and adolescents on the MMPI and MMPI-A.

Stein and Graham's (2005) most recent work using the MMPI-A was an extension of previous work in detecting defensiveness as cited in Stein & Graham (1999) and in detecting underreported substance use in adolescent offenders as cited in Stein & Graham (2001). They tested the efficacy of the L scale and substance abuse scales and found that the L scale was able to detect at least 75% of defensive profiles and 77% of honest profiles in substance-abusing and non-substance-abusing adolescent offenders. However, when instructed to fake good, it was difficult to detect defensiveness in both substance abusers and non-substance abusers utilizing the substance abuse scales. Therefore, in the detection of substance abuse, results implied that the L scale and substance abuse scales should be used in combination for the best chance at identifying defensive profiles.

Few studies to date have focused on assessment tools other than the MMPI and MMPI-2 with the exception of one. Wrobel et al. (1999) investigated response styles including defensiveness using the PIY validity scales. In a study designed to validate the PIY Defensiveness scale (designed to detect underreporting), Wrobel et al. (1999) found that inpatients who were tested upon admission then later retested with instructions to respond defensively on the PIY were able to produce profiles with markedly reduced clinical scales and subscale elevations. However, the PIY Defensiveness scale accurately detected faking good 90% of the time.

In summary, research suggests that adolescents can successfully fake good when instructed to respond defensively. Further, adolescents that are older and psychologically healthier appear to be better equipped to fake good. There is empirical support for the

validity scales on the MMPI, MMPI-A, and PIY to detect defensiveness. The L scale clearly appears to be most effective in detecting defensive responding on the MMPI and MMPI-A.

Need for Dissimulation Research with Preadolescents

Although there are some studies of adolescents, studies of dissimulation in preadolescents are absent from the literature. Prior to 1990, only a few general purpose, self-report measures with sound psychometric properties and clinical utility existed for adolescents and even less for younger children (Merrell, 2008). Self-report instruments for use with preadolescents that are currently available include the Youth Self-Report (YSR; Achenbach & Rescorla, 2001); the child and adolescent self-report forms from the Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004); the Children's Depression Inventory (CDI; Kovacs, 2003); and the Millon Pre-Adolescent Clinical Inventory (M-PACI; Millon, Tringone, Millon, & Grossman, 2005). Many self-report instruments commonly used in school psychology practice, such as the YSR and CDI, do not include specific scales to detect dissimulation. Of those instruments that contain validity scales designed to assess response styles, scales are limited or have not been adequately validated—as evidenced by the dearth of research in this area.

One measure that includes validity scales designed to assess dissimulation but lacks validation data on these scales is the BASC-2 Self-Report of Personality. The BASC-2 SRP (Reynolds & Kamphaus, 2004) is a comprehensive system of assessment instruments that include a self-report, called the Self-Report of Personality, for children ages 8 to 11. This instrument includes validity scales designed to detect malingering and

defensiveness, as well as irrelevant and random responding; however, descriptions of scale development are limited and information about validation procedures for the scales is absent from the manual (Rogers, 2008). The authors noted that two types of response formats (i.e., true/false and 4-point frequency scale) used on the BASC-2 SRP may be useful as another method to detect deliberate attempts to distort answers and to identify certain response sets.

The detection of children's dissimulation patterns is important due to the high stakes implications of assessment. School psychologists and other child-oriented psychologists are frequently called upon to provide recommendations based on their evaluations of children in various arenas including special education, juvenile court, divorce and custody hearings, child abuse allegations, and treatment planning. Further, they may be asked to provide recommendations regarding a child's ability to testify in investigations related to child welfare or domestic violence. Common referrals received by school psychologists are to determine the presence or absence of attention deficit hyperactivity disorder (ADHD) or internalizing disorders. According to the National Center for Health Statistics (2008), prevalence of ADHD has increased approximately 3% annually from 1997 to 2006. Externalizing symptoms (i.e., hyperactivity and impulsivity) associated with ADHD are more easily observable than those associated with inattention, which highlights the importance of subjective self-report in the evaluation process to adequately capture the full spectrum of symptomatology.

Prevalence rates of common internalizing disorders (e.g., depressive and anxiety disorders) range from approximately 2% -8% in normal samples of children (Anderson, Williams, McGee, & Silva, 1987; Costello, 1989). However, prevalence rates may be

underestimated because internalizing disorders are likely to be overlooked due to the private nature of associated symptoms (Reynolds, 2002), which, again, underscores the importance of self-report as a component of assessment. The utility of evaluation findings and recommendations depends upon a valid and comprehensive assessment, including reliable, accurate self-reports. Furthermore, the need for universal screening of emotional and behavioral problems in school settings is becoming increasingly popular as a means of early identification and prevention (New Freedom Commission on Mental Health, 2003; United States Public Health Service, 2000) and may include self-report tools.

Purpose of the Study

Although self-report measures are widely employed in the assessment of preadolescents, empirical data regarding dissimulation response styles in this age group is absent in the literature. To date, research has not been published with respect to the validity scales of the BASC-2 SRP, a preeminent omnibus self-report for young children (Merrell, 2008). This investigation examined the extent to which the BASC-2 SRP validity scales identified dissimulation response styles in children.

Specific research questions addressed in this study included the following. First, are the validity scales useful in signaling caution when interpreting the scores for children who are instructed to malinger or to "fake bad" on their responses? Second, are the validity scales useful in signaling caution when interpreting the scores for children who are instructed to respond defensively or to "fake good"? The null hypothesis is that there will be no difference in mean scores on the BASC-2 SRP validity scales generated by children instructed to malinger as well as mean scores generated by children instructed to respond defensively.

Our experimental hypotheses for each condition, as shown in Table 2, are as follows. We hypothesized that children provided with standard instructions and instructed to respond as a typical child on the BASC-2 SRP would generate average scores across all validity indicators and composite scores. We predicted that children instructed to malinger would generate high scores on the Emotional Symptoms Index, Internalizing Problems, and Inattention/Hyperactivity composites, a low score on the Personal Adjustment composite, and an elevated F scale. We predicted that children instructed to

Table 2

Predicted Validity Index Scores and Composite Scores Based on Experimental Condition

Condition	ESI	Int	Inatt/Hyp	PA	F	L	V	Resp Patt	Consist
Control	Avg	Avg	Avg	Avg	Avg	Avg	Avg	Avg	Avg
Malinger	High	High	High	Low	High	Avg	Avg	Avg	Avg
Defensive	Low	Low	Low	High	Low	High	Avg	Avg	Avg

respond defensively would generate low scores on the Emotional Symptoms Index, Internalizing Problems and Inattention/Hyperactivity composites, a high score on the Personal Adjustment composite, a low F scale, and an elevated L scale.

Previous research employed analyses to determine optimal cutoff scores to differentiate between dissimulation response styles and to discriminate between clinical/correctional and nonclinical/noncorrectional populations on self-report measures (Herkov et al. 1991; Lucio et al., 2002; Stein & Graham, 1999, 2001). The BASC-2 SRP

includes a range of caution indicators to interpret scores on the validity scales, which are useful in practice. Therefore, chi-square tests of independence were used to compare frequencies of validity scores falling within caution range to determine accuracy in predicting children's dissimulation styles.

Chapter 2: Method

Participants

Participants were 118 children recruited from nine general education classrooms from a university-affiliated elementary school in the MidSouth. In terms of gender, 58 girls and 60 boys voluntarily participated based on informed parental consent and child assent. Based on age and reading-level requirements of the measure used, children ranged in age from 8 to 11 years and were in the third ($n = 52$), fourth ($n = 21$), or fifth ($n = 45$) grade. Their mean age was 9.31 ($SD = 1.03$). Of the girls, 27.6% were 8 year olds ($n = 16$), 27.6% were 9 year olds ($n = 16$), 29.3% were 10 year olds ($n = 17$), and 15.5% were 11 year olds ($n = 9$). Of the boys, 26.7% were 8 year olds ($n = 16$), 31.7% were 9 year olds ($n = 19$), 28.3% were 10 year olds ($n = 17$), and 13.3% were 11 year olds ($n = 8$). Participation rates were 91% of third graders, 36% of fourth graders, and 82% of fifth graders. The overall participation rate was 69%.

Measures

All psychometric data were obtained from the BASC-2 SRP (Reynolds & Kamphaus, 2004) child form. The BASC-2 SRP is a self-report measure for children ages 8 to 11 containing items describing social and emotional characteristics and behaviors. The SRP consists of 139 items that are answered in two formats: true/false responses and ratings on a 4-point frequency scale.

Clinical composites. The SRP yields five composite scores, but only four were included in the analyses (i.e., Emotional Symptoms Index, Internalizing Problems, Inattention/Hyperactivity, and Personal Adjustment) based on their relevance to the proposed study. Each composite yields an age-based T-score with a mean of 50 and a

standard deviation of 10. The Emotional Symptoms Index provides a global indicator of serious emotional disturbance, particularly internalizing disorders. Six scales (i.e., Social Stress, Anxiety, Depression, Sense of Inadequacy, Self-Esteem, and Self-Reliance) compose the Emotional Symptoms Index. High scores on this scale indicate high disturbance, whereas low scores indicate low disturbance. The Internalizing Problems composite broadly captures the inner distress that is characteristic of internalizing disorders. Six scales (i.e., Atypicality, Locus of Control, Social Stress, Anxiety, Depression, and Sense of Inadequacy) compose the Internalizing Problems composite. High scores on this scale indicate greater levels of distress, whereas low scores indicate lesser levels of distress. Internal consistency estimates were .95 or greater and consistent between combined-gender and separate-gender groups at different age levels. The test–retest reliability estimate of .82 between test administrations with a 13-66 day interval was strong. The Inattention/Hyperactivity composite consists of the Attention Problems and Hyperactivity scales and is a good indicator of these symptoms. High scores on this scale indicate higher frequency of symptoms, whereas low scores indicate lower frequency of symptoms. The Personal Adjustment composite consists of the Relations with Parents, Interpersonal Relations, Self-Esteem, and Self-Reliance scales and is a reflection of interpersonal and intrapersonal functioning. High scores on this scale indicate healthier levels of adjustment, whereas low scores indicate poor levels of adjustment. These four BASC-2 SRP composites are supported by a wide array of validity evidence, including evidence based on content, internal relations, external relations with other measures of psychopathology and adaptive functioning, and group differences (Reynolds & Kamphaus, 2004).

Validity indexes. Validity indexes of the SRP include the F index, L index, V index, Response Pattern index, and Consistency index. The F index includes 15 items designed to detect excessive negativity in responses (e.g., I *never* succeed at anything I do; I *almost always* ignore people who are talking to me). Malingering may be suspected when respondents endorse an unusually high number of negative items, extreme severity of symptoms, or an unusually low number of positive items. An elevated F Index may also reflect a distress signal for respondents who genuinely present with severe emotional or behavioral disturbances. Alternatively, a high F index may also reflect poor comprehension or reading difficulties, random responding, or a problem following instructions. Profiles with raw scores on the F scale that fall between 0 and 3 are considered to be acceptable, those that fall between 4 and 6 should be interpreted with caution, and those that are 7 or more should be interpreted with extreme caution. The scale was developed by including items that were endorsed rarely (less than 3%) in the item-development samples. Reliability evidence for the scale is not reported in the BASC-2 manual.

The 13-item L index assesses the degree to which respondents attempt to portray themselves in a favorable light or "fake good." An elevated L index may reflect defensiveness, limited insight, poor reading skill, lack of comprehension, or random responding. The construction of this index was similar to that of the F index by including items that were endorsed rarely in the norm samples. L scale raw scores that fall between 0 and 9 are considered to be acceptable, those that fall between 10 and 12 signal caution, and scores 13 or above indicate extreme caution. Reliability evidence for the scale is not reported in the BASC-2 manual.

The V index includes 5 nonsensical items designed to detect carelessness in responding, poor comprehension of content, or uncooperativeness (e.g., I eat 25 apples every day). V index raw scores that fall between 0 and 2 are acceptable, those between 3 and 4 indicate caution, and those 5 or higher signify extreme caution. Scale development information is not provided in the BASC-2 manual, although authors reported that 3% of cases signifying extreme caution from the initial SRP forms collected during standardization were removed from the final sample used in the last stages of development of the norms. Reliability evidence for the scale is not reported in the BASC-2 manual.

The Response Pattern Index is designed to detect repetitive patterns of responding, which may indicate inattentiveness to item content. This index is computed by counting the number of times a response differs from the previous response. Low scores between 0 and 50 signal caution due to repetitive responding. High scores between 102 and 139 also signal caution due to frequent variation such as cyclical or alternating patterns. Those scores that fall between 51 and 101 are considered to be acceptable. Frequency distributions for the norm samples were reviewed in order to identify the extreme half-percentages at each end to create these ranges for score interpretation. Reliability evidence for the scale is not reported in the BASC-2 manual.

The Consistency index assesses the degree to which respondents answer like items in a similar manner. This index may also reflect inattention to items or misunderstanding of item content. Alternatively, the Consistency index could be an indication that the respondent changed their perspective during the administration. Scores between 0 and 16 are acceptable, those between 17 and 25 should be interpreted with

caution, and scores of 26 or more signal extreme caution. Using the item-development samples, the scale was constructed by (a) creating an item-correlation matrix in order to determine item pairs with the highest correlations and (b) selecting 20 item-pairs to contribute to this index. Reliability evidence for the scale is not reported in the BASC-2 manual.

Procedure

This investigation employed a between-subjects experimental simulation design with three conditions. Children who participated in this study were randomly assigned to a control group or to one of two simulation groups: (a) a malingering group to simulate behavioral and emotional symptoms, or (b) a defensive group to simulate minimization or denial of behavioral and emotional symptoms.

Recruitment. Letters of informed parental consent were sent home with all students in the third, fourth, and fifth grades at a university-affiliated elementary school in the MidSouth. To ensure receipt and maximize participation, letters were sent home on three separate occasions over a period of two weeks. Rosters were created for each classroom to include the names of all students with signed parental consent. Those students whose parents did not consent were dismissed from the classroom by their teacher and sent to the library. Next, child assent forms were passed out, a developmentally appropriate description of participation tasks and confidentiality was provided, and the voluntary nature of participation was explained.

Randomization. Each BASC-2 SRP protocol packet was labeled with a two-digit participant identification number (i.e., Participant #00001, Participant #00002, Participant #00003, etc.). Each label also included a unique six-digit number code to represent the

condition. Five-digit identification numbers and number codes were used in lieu of simple ordinal numbers (which may be associated with a subjective value or rating) in order to control for potential confounds. Research Randomizer software Copyright © 1997-2010 Geoffrey C. Urbaniak and Scott Plous (available free and online at http://www.randomizer.org/) was used to generate one set of 200 non-unique, unsorted numbers with a range from 1 to 3 (to correspond with the last digit in the number code for each condition). Numbers were generated 48 hours prior to data collection in order to minimize the time the primary investigator had knowledge of the random numbers generated and the administration of the BASC-2 SRP. Protocols were labeled and stacked based on the order in which unsorted numbers were generated. Protocols were distributed to children in the same manner in each classroom from the top of the stack and from left to right.

General administration. Test administration was conducted in groups by classroom and included nine classrooms. At least two adults were present for each administration. They included the classroom teacher and the primary investigator. The classroom teacher did not participate in the procedures but remained in the room to provide general supervision. In addition to the teacher and primary investigator, one research assistant was present in six of the nine classrooms to assist with administration. Research assistants were three graduate students studying school psychology that were trained in the study procedures.

The primary investigator instructed the children that packets would be passed out and that they were not to be opened until instructed to do so. Once each child had received a packet, the primary investigator instructed them to tear off the blank cover

sheet, to begin reading the standard instructions on the BASC-2 SRP, then to read the control or experimental group instructions on the next page of their packets. In each condition, children read the standard directions included on the BASC-2 SRP first, followed by specific instructions for one of two simulation conditions or the control group. Finally, a verbal reminder was delivered to the children as they completed the BASC-2 SRP at 5-minute intervals that consisted of the following statements, "Remember to pretend that you are another person and not to respond the way you normally would. You can look at the instruction pages if you need to. " One or two investigators circulated the classroom to monitor their progress and to clarify or repeat instructions, as needed. As the children finished, packets were collected by the investigators and maintained securely in one large envelope per classroom. Envelopes were sealed while data was transported.

Experimental conditions. Instructions for test administration in the experimental conditions were adapted from previous studies (Stein & Graham, 1999; Stein et al., 1995). Scripts were developed to ensure a similar number of words in each condition that ranged from 131 to 138 words. Readability was judged and calculated using the Flesch–Kincaid Grade Level Formula (Kincaid, Fishburne, Rogers, & Chissom, 1975) and ranged from 4.7 to 4.8.

Children in the malingering group were instructed as follows, "When you answer the questions, pretend that you are another person. Do not answer as you would normally. Answer the questions to make it seem like you have many behavior and emotional problems. Answer so it seems like you always get into trouble at school. When you answer, make it seem like you never pay attention when doing schoolwork. Make it seem

like you never stay seated and still when you should. Answer the questions to make it seem like you are always sad and nervous. Make it seem like you never get along with your classmates. Answer the questions to make it seem like you have many very serious behavior and emotional problems. Remember to pretend that you are another person. Remember not to answer as you would normally."

Children in the defensive group were instructed as follows, "When you answer the questions, pretend that you are another person. Do not answer as you would normally. Answer the questions to make it seem like you do not have serious behavior or emotional problems. Answer so it seems like you never get into trouble at school. When you answer, make it seem like you always pay attention when doing schoolwork. Make it seem like you always stay seated and still when you should. Answer the questions to make it seem like you are never sad and nervous. Make it seem like you always get along with your classmates. Answer the questions to make it seem like you have no behavior or emotional problems. Remember to pretend that you are another person. Remember not to answer as you would normally."

Children in the control group were provided with standard instructions in accordance with administration procedures for the BASC-2 SRP with an important exception. Children were not asked to provide information about themselves, but rather, to answer as they believed a typical child their age would respond. "When you answer the questions, pretend that you are another person. Do not answer as you would normally. Answer the questions to make it seem like you are a normal student in the class. Make it seem like you have no more behavior or emotional problems than anyone else. When you answer, make it seem like you pay attention as well as normal students when doing

schoolwork. Make it seem like you are no more sad or nervous than anyone else. Make it seem like you get along with your classmates as well as others. Answer the questions to make it seem like you are someone who is trying to be honest about how they think and feel and act. Remember to pretend that you are another person. Remember not to respond as you would normally."

Compliance Check

Upon completion of the BASC-2 SRP, all children were instructed to complete an eight-item integrity questionnaire (see Appendix A) to indicate how well they understood and complied with the instructions. Items required dichotomous yes or no responses and queried whether children were thinking about their own thoughts, feelings, and behaviors, which would not indicate compliance with the instructions, or the thoughts, feelings, and behaviors of the person described in the instructions for the experimental conditions, which would indicate compliance with the instructions. One general item targeted whether the children followed the directions. Four items targeted whether the children answered the questions about the person described in the instructions rather than themselves. Three items targeted whether the children answered the questions about themselves rather than the person described in the instructions.

Data Analysis

Comparisons of mean scores on BASC-2 SRP composites and validity indicators were conducted to test the hypotheses shown in Table 2. First, responses across conditions for each clinical composite were compared using one-way ANOVAs to determine if the differing instructions led to changes in the scores on the BASC-2 SRP that possess the most reliability and validity evidence supporting them. Next, responses

across conditions were compared using one-way ANOVAs for the validity indexes,

which have no reliability evidence and little validity evidence supporting their use. In

addition, an analysis targeting the validity scales was completed that examined the

frequency of "caution" values as yielded from the BASC-2 scoring software (see validity

indexes subsection in the Measures section). For this analysis, results indicating "some

caution" (or "caution-low" in the case of the Response Pattern Index) and "extreme

caution" (or "caution-high" in the case of the Response Pattern Index) were collapsed

into one category. For the F Index, values exceeding 3 indicated caution. For the L Index,

values exceeding 9 indicated caution. For the V Index, values exceeding 2 indicated

caution. For the Response Pattern Index, values below 51 and values exceeding 101

indicated caution. For the Consistency Index, values exceeding 16 indicated caution. An

a priori alpha level of .05 was used for all analyses. For all analyses in which a family of

analyses exceeded five, the Sidak–Bonferroni correction ($\alpha = .0057$) was used to control

for alpha inflation.

Chapter 3: Results

Demographic Characteristics

Demographic characteristics for children in each of the three conditions are presented in Table 3. Chi-square tests of independence revealed that there were no

Table 3

Demographic Characteristics of Children by Experimental Condition

Characteristic	Malinger $(N=43)$	Defensive $(N=42)$	Control $(N=33)$	$F(2, 115)$	χ^2
Age					
8 yrs.	12	13	7		
9 yrs.	13	12	10		
10 yrs.	10	13	11		
11 yrs.	8	4	5		
Condition *M*	9.33	9.19	9.42	.49	
Condition *SD*	1.09	0.99	1.00		
Grade					
3rd	20 (46.5%)	21 (50.0%)	11 (33.3%)		3.13
4th	9 (20.9%)	6 (14.3%)	6 (18.2%)		
5th	14 (32.6%)	15 (35.7%)	16 (48.5%)		
Gender					
Boys	19 (44.2%)	28 (66.7%)	13 (39.4%)		6.70*
Girls	24 (55.8%)	14 (33.3%)	20 (60.6%)		

Note. *p < .05

27

significant differences in the frequency of children at different grade levels across conditions, $\chi^2(4, N = 118) = 3.13, p = .536$. One-way analysis of variance (ANOVA) revealed no significant differences in children's ages across conditions $F(2, 115) = .49, p = .614$. However, there was a significant difference between groups for gender, $\chi^2(2, N = 118) = 6.70, p = .035$. There were significantly more boys than girls in the defensive condition and significantly more girls than boys in the control condition.

Gender Effects

In order to determine if the significant difference in frequency of boys and girls across conditions had an effect on results, independent *t*-tests were used to test for gender effects on clinical composite and validity index scores across conditions. No significant differences were found between boys' and girls' mean scores on the clinical composite and validity index scores. Independent *t*-tests were also used to test for gender effects within conditions. The Sidak–Bonferroni correction ($\alpha = .0019$) was used to control for alpha inflation in the analyses within conditions. It was only in the control condition that boys ($M = 71.33, SD = 9.10$) produced lower scores than girls ($M = 83.86, SD = 8.38$) on the Response Pattern Index; $t(39) = -4.29, p < .001$. Thus, it does not appear that gender effects skewed the results from the analysis of the composite scores and validity indexes across conditions and within conditions due to the disproportionate sampling of boys and girls across conditions.

Compliance Check

The compliance check included eight items requiring a response of yes or no. A high number of items marked suggesting compliance indicated that the participant had good comprehension of the experimental instructions, whereas a lower number of items

marked as suggesting compliance indicated weaker comprehension of the experimental

instructions. Of the total sample, one participant in the defensive condition did not

complete the compliance check. Of the 117 children who completed the compliance

check (see Table 4), 16.1% marked all eight items suggesting compliance, 15.3% marked

Table 4

Compliance Frequency and Percentages across Conditions and by Experimental Condition

Item measuring compliance and response indicating compliance	Across Conditions	Condition			χ^2 or F
		Malinger	Defensive	Control	
1. Did you understand the instructions? (Y)	109 92.4%	40 93.0%	38 92.7%	31 93.9%	0.05
2. When you answered, were you thinking about how the person described in the instructions would think or feel? (Y)	96 82.1%	34 79.1%	36 87.8%	26 78.8%	1.42
3. When you answered, were you thinking about how the person described in the instructions would answer? (Y)	102 87.2%	36 83.7%	37 90.2%	29 87.9%	0.82
4. When you answered the questions, were you thinking about how you think or feel or behave? (N)	66 42.6%	20 52.4%	24 40.0%	22 33.3%	2.91
5. When you answered the questions, were you thinking about your own answers to the questions? (N)	65 44.4%	22 48.8%	22 46.3%	21 36.4%	1.27

(table continues)

29

Table 4 (Continued)

Compliance Frequency and Percentages across Conditions and by Experimental Condition

Item measuring compliance and response indicating compliance	Across Conditions	Condition			χ^2 or F
		Malinger	Defensive	Control	
6. Do you think you answered the questions like the person described in the instructions would answer? (Y)	94 80.3%	35 81.4%	34 82.9%	25 75.8%	0.64
7. Do you think that your answers describe how you think, feel or behave? (N)	70 38.6%	20 52.4%	24 38.5%	26 21.2%	7.58*
8. Do you think your answers describe how a pretend person thinks, feels or behaves? (Y)	94 81.0%	30 71.4%	36 87.8%	28 84.8%	4.06
Overall Compliance *M* (Number of items indicating compliance endorsed)	5.46	5.58	5.61	5.12	0.86
Overall Compliance *SD*	1.76	2.08	1.62	1.45	

Note. *$p < .05$

seven items suggesting compliance, 12.7% marked six items suggesting compliance,

29.7% marked five items suggesting compliance, 12.7% marked half of the items

suggesting compliance, and 12.8% marked three or fewer items suggesting compliance.

Overall, more than 50% of children across all three conditions indicated

compliance on the item that globally assessed whether they followed the directions as

well as on all four items keyed in a positive direction that assessed whether they thought

about the person in the instructions rather than themselves while completing the BASC-2

SRP. However fewer than 50% of children across conditions indicated compliance on all three items keyed in a negative direction that assessed whether they thought about themselves rather than the person in the instructions while completing the BASC-2 SRP.

As evident in Table 4, chi-square tests of independence revealed no significant differences across conditions in the frequency with which children indicated compliance on the first item that globally targeted whether they followed the directions suggesting that children understood equally and complied equally with the instructions across conditions. One-way analysis of variance (ANOVA) revealed no significant differences in the mean number of items marked indicating compliance across conditions $F(2, 114) =$ .86, $p = .428$ also suggesting that children in each condition (malingering $M = 5.58$, $SD = 2.08$; defensive $M = 5.61$, $SD = 1.62$; control $M = 5.12$, $SD = 1.45$) understood equally and complied equally with instructions.

No significant differences were found across conditions in the frequency with which children indicated compliance on the four items keyed in a positive direction that assessed whether they thought about the person in the instructions rather than themselves while completing the BASC-2 SRP. However, on the three items keyed in a negative direction that assessed whether they thought about themselves rather than the person in the instructions while completing the BASC-2 SRP, proportionately fewer children in the control condition indicated compliance and proportionately more in the malingering condition indicated compliance.

Clinical Composites

Table 5 presents means, standard deviations, and skewness and kurtosis statistics for the clinical composite scores by condition. Prior to data analysis, the Emotional

Table 5

Means and Standard Deviations of Clinical Composite Scores by Experimental Condition

	Malinger	Defensive	Control	ANOVA and Tukey HSD Results
Externalizing Symptoms				*F(2*, 114) = 53.23*
M	71.40	43.76	48.42	M > D and C
SD	17.33	9.32	10.25	
Skewness	-.20	1.46	1.42	
Kurtosis	-1.20	2.12	2.49	
Internalizing Problems				*F(2*, 114) = 49.08*
M	67.56	44.10	49.67	M > D and C
SD	14.24	8.67	9.74	
Skewness	-.08	1.42	1.13	
Kurtosis	-1.23	1.74	1.54	
Inattention/Hyperactivity				*F(2*, 115) = 44.50*
M	69.35	43.45	49.18	M > D and C
SD	17.13	10.92	9.47	
Skewness	-.46	2.00	.48	
Kurtosis	-1.04	3.73	-.45	
Personal Adjustment				*F(2*, 114) = 50.76*
M	27.77	53.88	51.03	M < D and C
SD	17.36	8.63	10.28	
Skewness	.56	-.75	-1.02	
Kurtosis	-.96	-.08	.30	

Note. * $p < .05$. M = Malingering; D = Defensive; C = Control.

Symptoms Index, Internalizing Problems, Inattention/Hyperactivity, and Personal

Adjustment variables were screened (within conditions) for missing data, the presence of

univariate outliers, normality of the distributions, and equality of variances using various

SPSS tests of the assumptions for ANOVA. Cases with *z*-scores in excess of 3.29 were

identified as potential outliers, as recommended by Tabachnick and Fidell (2007). Only a

case in the defensive condition was identified as a univariate outlier based on an

extremely high *z*-score on the Inattention/Hyperactivity Index, but this case was not

omitted. As shown in Table 5, the distributions for the Internalizing Problems and

Personal Adjustment composites under all three conditions were rather normal in shape.

However, statistical analyses of the Emotional Symptoms Index composite under the

defensive and control conditions and the Inattention/Hyperactivity composite under the

defensive condition suggested non-normal distributions in terms of kurtosis.

All clinical composite scores failed Levene's test for equality of variances across

conditions. For these measures, Welch's (1951) *F*' test was used to compute univariate

analyses of variance for each clinical composite across all groups. This test does not

require equal sample sizes or homogeneity of variance. The Robust Tests of Equality of

Means indicated that, for each clinical composite, there were statistically significant

differences between groups. Considering that Levene's tests may be too sensitive to

equality of variance, that results of the ANOVA and F' tests were similar, and that

ANOVA is reasonably robust against violations of some of these assumptions, we

reported ANOVA results with post-hoc tests.

Emotional Symptoms Index. As shown in Table 5, a one-way between subjects

ANOVA was conducted to compare the effect of instructing children to simulate their

responses on the BASC-2 SRP on the Emotional Symptoms Index in the malingering,

defensive, and control conditions. As hypothesized, there was a significant effect of type

of instructions on the Emotional Symptoms Index for the three conditions, $F(2, 114) =$

53.23, $p < .001$. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for the malingering condition ($M = 71.40$, $SD = 17.33$) was significantly higher than the control condition ($M = 48.42$, $SD = 10.25$) and the defensive condition ($M = 43.76$, $SD = 9.32$), $p < .05$. Contrary to expectations, the defensive condition ($M = 43.76$, $SD = 9.32$) was not significantly lower than the control condition ($M = 48.42$, $SD = 10.25$). The assumption of homogeneity of variance was violated; therefore, the Welch F-ratio is reported. Consistent with ANOVA, the Welch version of the F-ratio indicated that there was a significant effect of type of instructions on the Emotional Symptoms Index, $F(2, 72.49) = 41.91$, $p < .001$.

Internalizing Problems. To test the hypothesis that instructing children in the malingering, defensive, and control conditions to simulate their responses on the BASC-2 SRP would selectively affect their scores on the Internalizing Problems composite, a one-way between subjects ANOVA was conducted. As evident in Table 5, there was a significant effect of type of instructions on the Internalizing Problems Composite for the three conditions $F(2, 114) = 49.08$, $p < .001$. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for the malingering condition ($M = 67.56$, $SD = 14.24$) was significantly higher than the control condition ($M = 49.67$, $SD = 9.74$) and the defensive condition ($M = 44.10$, $SD = 8.67$), $p < .05$. However, the defensive condition ($M = 44.10$, $SD = 8.67$) was not significantly lower than the control condition ($M = 49.67$, $SD = 9.74$). The assumption of homogeneity of variance was violated; therefore, the Welch F-ratio was calculated and also indicated a significant effect of type of instructions on the Internalizing Problems Composite, $F(2, 72.77) = 41.76$, $p < .001$.

Inattention/Hyperactivity. A one-way between subjects ANOVA was conducted to compare the effect of instructing children to simulate their responses on the BASC-2 SRP on the Inattention/Hyperactivity Composite in the malingering, defensive, and control conditions. As shown in Table 5, there was a significant effect of type of instructions on the Inattention/Hyperactivity Composite for the three conditions $F(2, 115) = 44.50, p < .001$. As hypothesized, post-hoc comparisons using the Tukey HSD test indicated that the mean score for the malingering condition ($M = 69.35, SD = 17.13$) was significantly higher than the control condition ($M = 49.18, SD = 9.47$) and the defensive condition ($M = 43.45, SD = 10.92$), $p < .05$. Contrary to expectations, the defensive condition ($M = 43.45, SD = 10.92$) was not significantly lower than the control condition ($M = 49.18, SD = 9.47$). The assumption of homogeneity of variance was violated; therefore, the Welch F-ratio is reported. Consistent with ANOVA, the Welch version of the F-ratio indicated that there was a significant effect of type of instructions on the Inattention/Hyperactivity Composite, $F(2, 75.37) = 34.77, p < .001$.

Personal Adjustment. To test the hypothesis that instructing children to simulate their responses on the BASC-2 SRP in the malingering, defensive, and control condition would selectively impact their scores on the Personal Adjustment composite, a one-way between subjects ANOVA was conducted. As evident in Table 5, there was a significant effect of type of instructions on the Personal Adjustment Composite for the three conditions $F(2, 114) = 50.76, p < .001$. Post-hoc comparisons using the Tukey HSD test indicated that, indeed, the mean score for the malingering condition ($M = 27.77, SD = 17.36$) was significantly lower than the control condition ($M = 51.03, SD = 10.28$) and the defensive condition ($M = 53.88, SD = 8.63$), $p < .05$. However, the defensive condition

($M = 53.88$, $SD = 8.63$) was not significantly higher than the control condition ($M = 51.03$, $SD = 10.28$). The assumption of homogeneity of variance was violated; therefore, the Welch F-ratio was calculated and also indicated a significant effect of type of instructions on the Personal Adjustment Composite, $F(2, 71.10) = 38.95$, $p < .001$.

In summary, these findings revealed that children in the malingering condition were consistently able to successfully fake bad as indicated by mean scores on all four clinical composites that fell within the At-Risk or Clinically Significant ranges. Also, children in the defensive and control conditions produced scores that indicated little psychopathology, but there were no significant differences between these groups.

Validity Indexes

Table 6 presents means, standard deviations, and skewness and kurtosis statistics for the clinical composite scores by condition. Prior to data analysis, the F Index, L Index, V Index, Response Pattern Index, and Consistency Index variables were screened for missing data, the presence of univariate outliers, normality of the distributions, and equality of variances using various SPSS tests of the assumptions for ANOVA. Cases with z-scores in excess of 3.29 were identified as potential outliers, as recommended by Tabachnick and Fidell (2007). Two cases in the defensive condition were identified as univariate outliers based on their extremely high z-scores on the F Index and V Index, respectively. The distributions for the L Index, V Index, Response Pattern Index, and Consistency Index under all three conditions were rather normal in shape. However, statistical analyses of the F Index distributions under the defensive and control conditions suggested that scores were distributed non-normally in terms of kurtosis. Two measures, the F Index and V Index, failed Levene's test for equality of variances. For these

Table 6

Means and Standard Deviations of Validity Index Scores by Experimental Condition

	Malinger	Defensive	Control	
F index				$F(2, 114) = 35.47*$
M	5.86	0.61	1.00	M > D and C
SD	4.84	1.20	1.68	
Skewness	.27	2.27	1.95	
Kurtosis	-1.21	4.85	3.30	
L index				$F(2, 114) = 21.55*$
M	3.19	7.88	5.58	M < D and C
SD	3.21	3.45	3.13	
Skewness	1.224	-.475	.186	
Kurtosis	.88	-.65	-.55	
V index				$F(2, 114) = 3.50$
M	2.53	1.20	1.42	
SD	3.14	1.81	2.17	
Skewness	1.20	1.86	1.34	
Kurtosis	.46	3.92	.76	
Response Pattern				$F(2, 114) = 4.94$
M	82.40	75.61	83.79	
SD	13.16	10.62	13.03	
Skewness	-.11	.09	.16	
Kurtosis	.53	-.57	.09	
Consistency				$F(2, 114) = 7.48*$
M	11.14	6.37	9.52	
SD	5.68	4.67	6.86	
Skewness	.47	.33	.91	
Kurtosis	-.19	-1.02	.25	

Note. $* p < .05$. M = Malingering; D = Defensive; C = Control.

measures, Welch's (1951) *F'* test was used to compute univariate analyses of variance for each factor across all groups. The Robust Tests of Equality of Means indicated that, for the F Index, there were statistically significant differences between groups. Considering that Levene's tests may be too sensitive to equality of variance, that results of the ANOVA and F' tests were similar, and that ANOVA is reasonably robust against violations of some of these assumptions, we reported ANOVA results with post-hoc tests.

$\quad$ *F Index*. A one-way between subjects ANOVA was conducted to compare the effect of instructing children to simulate their responses on the BASC-2 SRP on the F Index in the malingering, defensive, and control conditions. As evident in Table 6, there was a significant effect of type of instructions on the F Index for the three conditions $F(2, 114) = 35.47$, $p < .001$. As hypothesized, post-hoc comparisons using the Tukey HSD test indicated that the mean score for the malingering condition ($M = 5.86$, $SD = 4.84$) was significantly higher than the control condition ($M = 1.00$, $SD = 1.68$) and the defensive condition ($M = 0.61$, $SD = 1.20$), $p > .05$. However, the defensive condition ($M = 0.61$, $SD = 1.20$) did not differ significantly from the control condition ($M = 1.00$, $SD = 1.68$). The assumption of homogeneity of variance was violated; therefore, the Welch *F*-ratio is reported. Consistent with ANOVA, the Welch version of the *F*-ratio indicated that there was a significant effect of type of instructions on the F Index, $F(2, 65.68) = 23.57$, $p < .001$.

$\quad$ *L Index.* To test the hypothesis that instructing children to simulate their responses on the BASC-2 SRP in the malingering, defensive, and control conditions would selectively affect their scores on the L Index, a one-way between subjects ANOVA was conducted. As shown in Table 6, there was a significant effect of type of instructions on

the L Index for the three conditions $F(2, 114) = 21.55$, $p < .001$. Post-hoc comparisons

using the Tukey HSD test indicated that the mean score for the malingering condition (M

$= 3.19$, $SD = 3.21$) was significantly lower than the control condition ($M = 5.58$, $SD =$

3.13) and the defensive condition ($M = 7.88$, $SD = 3.45$), $p < .05$. Also, the mean score

for the control condition ($M = 5.58$, $SD = 3.13$) was significantly lower than the defensive

condition ($M = 7.88$, $SD = 3.45$), $p < .05$.

V Index. As evident in Table 6, a one-way between subjects ANOVA was

conducted to compare the effect of instructing children to simulate their responses on the

BASC-2 SRP on the V Index in the malingering, defensive, and control conditions. As

hypothesized, there was not a significant effect of type of instructions on the V Index for

the three conditions $F(2, 114) = 3.50$, $p = .033$. The assumption of homogeneity of

variance was violated; therefore, the Welch F-ratio was calculated and also indicated that

there was not a significant effect of type of instructions on the V Index, $F(2, 71.73) =$

2.92, $p = .06$.

Response Pattern Index. A one-way ANOVA was conducted to compare the

effect of instructing children to simulate their responses on the BASC-2 SRP on the

Response Pattern Index in the malingering, defensive, and control conditions. As

hypothesized, there was no significant effect of type of instructions on the Response

Pattern Index for the three conditions $F(2, 114) = 4.94$, $p = .009$.

Consistency Index. In a similar way, a one-way between subjects ANOVA was

conducted to compare the effect of instructing children to simulate their responses on the

BASC-2 SRP on the Consistency Index in the malingering, defensive, and control

conditions. As shown in Table 6, there was a significant effect of type of instructions on

the Consistency Index for the three conditions $F(2, 114) = 7.48$, $p = .001$. Post-hoc

comparisons using the Tukey HSD test indicated that the mean score for the malingering

condition ($M = 11.14$, $SD = 5.68$) was significantly higher than the defensive condition

($M = 6.37$, $SD = 4.67$), $p < .05$. However, the malingering condition ($M = 11.14$, $SD = 5.68$) did not differ significantly from the control condition ($M = 9.52$, $SD = 6.86$) and the

control condition ($M = 9.52$, $SD = 6.86$) did not differ significantly from the defensive

condition ($M = 6.37$, $SD = 4.67$).

In summary, these findings suggest that the F Index and L Index consistently

detect faking bad. Specifically, a high F Index and a low L Index seem to be good

indicators of malingering. However, the L Index did not appear to be as sensitive to

defensive responding.

Analysis by Caution Category. The validity scales were targeted for analysis by

examining the frequency of "caution" values as yielded from the BASC-2 scoring

software. For this analysis, results indicating "some caution" (or "caution-low" in the

case of the Response Pattern Index) and "extreme caution" (or "caution-high" in the case

of the Response Pattern Index) were collapsed into one category. As shown in Table 7,

chi-square tests of independence revealed that there were no significant differences in the

frequency of scores in caution range for the V Index and the Response Pattern Index

across conditions. However, there were significantly more scores in caution range on the

F Index in the malingering condition, $\chi^2(2, N = 117) = 40.45$, $p = .000$ than in the

defensive and control conditions. There were also significantly more scores in caution

range on the L Index scores in the defensive condition, $\chi^2(2, N = 117) = 20.25$, $p = .000$,

Table 7

Percentage of Children Reporting Caution Levels on Validity Indicators

	Malinger ($N = 43$)	Defensive ($N = 41$)	Control ($N = 33$)	χ^2
Validity Indicator				
F Index	26 (60.5%)	2 (4.9%)	3 (9.1%)	40.45*
L Index	1 (2.3%)	16 (39.0%)	4 (12.1%)	20.25*
V Index	15 (34.9%)	8 (19.5%)	10 (30.3%)	
Response Pattern Index	4 (9.3%)	0 (0.0%)	2 (6.1%)	
Consistency Index	7 (16.3%)	0 (0.0%)	6 (18.2%)	7.96*

Note. *$p < .05$

defensive condition, $\chi^2(2, N = 117) = 7.96$, $p = .019$, yielded significantly fewer scores in

caution range than in the malingering and control conditions.

Chapter 4: Discussion

This investigation provided empirical evidence regarding dissimulation response styles in preadolescents that, until now, has been absent in the literature. The current study is the first to examine the extent to which the BASC-2 SRP validity scales detected dissimulation in children. Our findings indicate that the F Index and the L Index are the validity scales that are most sensitive in detecting children who were instructed to malinger. Specifically, children that generated high F Index scores and low L Index scores were most likely to be identified as having malingered or faked bad. Further, we were able to detect 61% of children who were instructed to malinger and 39% of children who were instructed to respond defensively by analyzing caution levels for the F Index and L Index, respectively.

Clinical Composites

As we hypothesized, children who were instructed to malinger generated high scores (within the Clinically Significant or At-Risk range) on the Emotional Symptoms Index, Internalizing Problems index, and Inattention/Hyperactivity index when compared to children in the defensive and control conditions who generated scores within the Average range. We also predicted that children who were instructed to malinger would generate lower Personal Adjustment scores than children who were instructed to respond defensively and children in the control condition. Indeed, our findings suggested that children in the malingering condition yielded scores within the Clinically Significant range (on average), whereas those in the defensive and control conditions yielded scores within the Average range on the Personal Adjustment index. These results support the idea that "typical" children are quite capable of faking bad and that the most reliable and

valid scores from the BASC-2 SRP, the clinical composites, are significantly elevated when children malinger. These results are consistent with Archer et al. (1987) who also found that, when adolescents were instructed to malinger, there were marked elevations within the clinical range on the clinical scales of the MMPI.

Contrary to our expectations, children who were instructed to respond defensively did not generate significantly lower scores on the clinical composites than children in the control conditions. Our results are similar to two studies (Archer, 1992; Baer et al., 1998) that reported mean scores within the Average range on the Clinical scales of the MMPI-A for adolescents who were instructed to fake good. One explanation for our results may be potential floor effects on the BASC-2 SRP test items comprising the clinical composites that affect their sensitivity in distinguishing between children in the control and defensive conditions. Additionally, a challenge in detecting defensiveness in children may be that developmental and maturational differences impact their ability to fully comprehend the distinctions between normal and abnormal thoughts, behaviors, and feelings. Archer (1987) found that younger adolescents were less effective than older adolescents at simulating normalcy. Alternatively, our results may indicate that children who were instructed to respond defensively were more successful than expected in simulating normalcy. Children in the defensive condition may have followed simulation instructions so well that they produced mean scores that were no different and no more deflated than those in the control condition. Thus, their scores did not signal caution and children who were instructed to respond defensively escaped detection. It is also notable that our sample consisted of typical children who may be better at simulating normalcy than

43

atypical children in clinical/correctional samples in previous investigations (Archer et al., 1987; Herkov et al., 1991; Stein & Graham, 1999, 2005).

Validity Indexes

As we hypothesized, children who were instructed to malinger generated significantly higher mean scores within the At-Risk and Clinically Significant ranges on the F Index when compared to children in the defensive and control conditions. Furthermore, 61% of children in the malingering condition produced scores on the F Index that signaled caution. Children in the malingering condition also produced scores that signaled caution significantly more frequently than children in the defensive and control conditions. This evidence suggests that the F Index performs well as a validity indicator for faking bad in children. Similarly, Archer (1987) found adolescents who were instructed to malinger generated marked elevations on the F scale of the MMPI, which was designed to detect highly unusual response patterns, including exaggeration of symptoms and problems. Although the F Index of the BASC-2 SRP and the F scale of the MMPI are not the same, they were both designed to detect faking bad.

We predicted that children who were instructed to respond defensively would generate higher scores on the L Index. Although children in the defensive condition yielded higher scores than children in the malingering condition, the mean score did not reach the caution range. However, 39% of children in the defensive condition produced scores on the L Index that signaled caution. Even though children in the defensive condition escaped detection in terms of mean score comparisons on the clinical composites and validity indexes scores, their scores signaled caution more frequently than those in the malingering and control conditions. This finding suggests that children who are instructed

to respond defensively and children who are instructed to respond normally may be unlikely to produce significantly different scores on the L Index.

Children who were instructed to malinger produced significantly lower scores on the L Index than children in the defensive and control groups. Although the L Index did not appear to perform as well as expected as an indicator of defensiveness, our findings suggest that a low score on the L Index may be a good indicator of malingering, particularly when there is also a high score on the F Index. Several studies demonstrated that profile analysis including the evaluation of the infrequency index (similar to the F Index) in relation to the lie scale score (such as the L Index) is a useful indicator of malingering (Lucio et al., 2002; Stein & Graham, 1999; Stein et al., 1995). Future research employing profile analysis of the F Index and L Index on the BASC-2 SRP may be warranted to determine if accuracy rates increase in the detection of malingering.

As hypothesized, children in the control and defensive conditions generated mean scores within the acceptable range on the V Index but children in the malingering condition generated a higher mean score approaching caution range, though the difference was not statistically significant. Similarly, the analysis by caution category identified proportionately more children in the malingering condition who produced V Index scores within the caution range than in the defensive and control conditions but, again, this difference did not reach statistical significance. In retrospect, it seems plausible that the V Index, which is similar in content to the F Index and includes nonsense items designed to detect carelessness in responding, failure to comprehend items or uncooperativeness, could be elevated for children who are instructed to malinger. The items comprising the V Index were endorsed even less frequently than

items comprising the F Index during standardization and include even more unusual content. Perhaps, typical kids are better at faking bad "realistically" by avoiding the overreporting of obviously absurd symptoms. Furthermore, this raises the question of whether "normal" children produce different scores in comparison to atypical children in clinical/correctional samples, which may warrant further study.

As hypothesized, children in the control and defensive conditions also generated scores within the acceptable range on the Response Pattern Index and Consistency Indexes. We found a gender difference in that boys generated significantly higher scores than girls on the Response Pattern Index; however, mean scores fell within the acceptable range. This suggests that there was no difference in the frequency with which Response Pattern Index scores signaled caution for boys when compared to girls. Children who were instructed to malinger generated significantly higher scores on the Consistency Index than those in the defensive and control conditions. However, this finding is not particularly meaningful since the mean score fell within the acceptable range. Children who were instructed to malinger may have changed their perspective while completing the BASC-2 due to confusing their instructed response style with their typical response style.

Limitations and Future Research

Our findings are subject to at least three limitations. First, our sample consisted of typical children recruited from general education classrooms. The absence of data from a clinical/correctional group that previously completed a related scale to provide evidence of "true" dissimulation limits our ability to make inferences from our results. It is

recommended that future work include a clinical or correctional sample for comparison purposes.

Another potential source of weakness in this study was that 25% of the children endorsed only half or fewer items to indicate compliance on the integrity questionnaire. They appeared to have more difficulty on the three items keyed in a negative direction that assessed whether they thought about themselves rather than the person in the instructions while completing the BASC-2 SRP. Proportionately fewer children in the control condition indicated compliance and proportionately more in the malingering condition indicated compliance on the items keyed in a negative direction. In future research, it would be interesting to assess compliance using an alternative method (e.g., Likert scale) that does not require children to shift their frame of reference when responding.

Thirdly, simulation instructions for the malingering and defensive conditions may have been too general in comparison to those used in previous studies. Experimental instructions in prior research with adolescents mentioned specific problems, described motives, referred to gaining attention or seeking adult support, or offered incentives. In consideration that our sample is the first to study dissimulation with a younger age group, instructions were written for ease of readability and comprehension. However, future research may wish to extend this work by including more detail in simulation instructions by referencing specific behaviors associated with common childhood disorders (e.g, ADHD, depression, anxiety) or by providing specific scenarios to elicit particular response styles.

Implications for Practice

This study offers some initial data on the validity scales of the BASC-2 SRP, which is widely used in educational and clinical settings for screening and diagnostic purposes. Our findings provided support that the F Index assesses the degree to which children respond in an overtly negative fashion and that the L Index assesses the degree to which children respond to present themselves in an unrealistically favorable light, as each scale was designed to do. Overall, scores yielded on the Response Pattern fell within the Acceptable range and visual inspection of protocols did not suggest significantly repetitive or cyclical patterns in responding. Current findings seem to support that this scale is, indeed, an indicator of attentiveness to item content. Similarly, the children in our sample generally produced scores on the Consistency Index that fell within the acceptable range providing support that this scale is also an indicator of attentiveness to item content as it was designed to be. Further work needs to be done to provide reliability data on the validity scales.

Our findings regarding the frequency with which validity indicators signaled caution yielded compelling results considering how the BASC-2 SRP is used in practice. Significantly more children in the malingering condition generated F Index scores that signaled caution than either the defensive or control groups. Further, there were significantly more children in the defensive group that generated L Index scores that signaled caution than either the malingering or control group. Based on our findings, we agree with Reynolds and Kamphaus (2004) that it is important to assess validity prior to any interpretation of scores on the BASC-2 SRP. We support an initial visual inspection of the protocol to scan for missing items, repetitive responding, or cyclical patterns of

responding followed by review of the validity indicators and interpretation of their caution ranges with specific attention to the F Index and L Index. We found that an elevated F Index and a low L Index appear to be effective in detecting malingering in children that were instructed to fake bad. It is suggested that the L Index caution range be reviewed to determine the possibility of defensive responding based on the greater frequencies that children who were instructed to respond defensively were detected in our study. In turning to the clinical composites, we posit that elevations across several of these scores combined with a high F index, a low Index, or both may be strong indicators of malingering.

According to the BASC-2 Manual, the SRP child form and adolescent form are similar in scales, structure, and content. However, the child form contains 40 fewer items than the adolescent form. It is worthwhile to consider whether more items on the child form would enhance its psychometric properties, including the accuracy of the validity scales in detecting dissimulation response styles.

In conclusion, this study set out to determine the extent to which the validity scales of the BASC-2 SRP identified dissimulation response styles in children. Our work makes noteworthy contributions to the existing literature on dissimulation as the first study investigating the BASC-2 in this capacity and extending this work to preadolescents. It is now possible to state that, based on our results, children who were instructed to malinger were able to successfully do so. However, their faking bad response style did not escape detection on the F Index and L Index. In contrast, children who were instructed to respond defensively were not as skilled at faking good but the L Index performed adequately in identifying their faking good response style. These

findings provide a starting point for continued study of children's dissimulation response styles and underscore the need for validity data in evaluating the performance of validity scales on self-report measures.

References

Achenbach, T. M., & Edelbrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs for the Society for Research in Child Development*, *46*, 1-82.

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

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children. *Archives of General Psychiatry*, *44*, 69-76.

Archer, R. P. (1992). *MMPI-A: Assessing Adolescent Psychopathology*. Hillsdale, NJ: Erlbaum.

Archer, R. P., Gordon, R. A., & Kirchner, F. H. (1987). MMPI response-set characteristics among adolescents. *Journal of Personality Assessment*, *51*, 505-516.

Archer, R. P., Maruish, M., Imhof, E. A., & Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology: Research and Practice*, *22*, 247-252.

Baer, R. A., Ballenger, J., & Kroll, L. S. (1998). Detection of underreporting on the MMPI-A in clinical and community samples. *Journal of Personality Assessment*, *71*, 98-113.

Behar, L. (1977). The Preschool Behavior Questionnaire. *Journal of Abnormal Child Psychology*, *5*, 265-275.

Bussey, K. (1992). Children's lying and truthfulness: Implications for children's testimony. In S. J. Ceci, M. D. Leichtman, & M. E. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 89-109). Hillsdale, NJ: Erlbaum.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic 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., Tellegen, A., Ben-Porath, Y. S., et al. (1992). *Minnesota Multiphasic Personality Inventory – Adolescent (MMPI-A): Manual for administration, scoring and interpretation.* Minneapolis: University of Minnesota Press.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development*, *60*, 1263-1277.

Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*, 836-841.

Ferguson, L. R., Partyka, L. B., & Lester, B. M. (1974). Patterns of parent perception differentiating clinic from nonclinic children. *Journal of Abnormal Child Psychology*, *2*, 169-181.

Gough, H. G. (1950). The F minus K dissimulation index for the MMPI. *Journal of Consulting Psychology*, *14*, 408-413.

Graham, J. R. (2000). *MMPI-2: Assessing personality and psychopathology,* (3rd ed.). New York: Oxford University Press.

Hathaway, S. R., & McKinley, J. C. (1940). A Multiphasic Personality Schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology*, *10*, 249-254.

Herkov, M. J., Archer, R. P., & Gordon, R. A. (1991). MMPI response sets among adolescents: An evaluation of the limitations of the subtle-obvious scales. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 424-426.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease formula) for Navy Enlisted Personnel. Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis.

Kovacs, M. (2003). *Children's Depression Inventory*. North Tonawanda, NY: Multi-Health Systems, Inc.

Lewis, M. (1993). The development of deception. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 90-105). New York: Guilford Press.

Lucio, E. (1998). *Spanish version of the Minnesota Multiphasic Personality Inventory: MMPI-A for Mexico*. Mexico City, Mexico: El Manual Moderno.

Lucio, E., Durán, C., Graham, J. R., & Ben-Porath, Y. S. (2002). Identifying faking bad on the Minnesota Multiphasic Personality Inventory-Adolescent with Mexican adolescents. *Assessment, 9*, 62-69.

McCann, J. T. (1998). *Malingering and deception in adolescents: assessing credibility in clinical and forensic settings*. Washington, DC: American Psychological Association.

Merrell, K. W. (2008). *Behavioral, social and emotional assessment of children and adolescents* (3rd ed.). New York: Lawrence Erlbaum Associates.

Millon, T., Tringone, R., Millon, C., & Grossman, S. (2005). *Millon Pre-Adolescent Clinical Inventory*. San Antonio, TX: Pearson.

National Center for Health Statistics. (2008). *Diagnosed attention deficit hyperactivity disorder and learning disability: United States 2004-2006*. (DHHS Publication No. PHS 2008-1565). Washington, DC: Government Printing Office.

New Freedom Commission on Mental Health. (2003). *Achieving the promise: Transforming mental health care in America. Final Report*. (DHHS Publication No. SMA-03-3832). Rockville, MD: Author.

Reynolds, W. M. (2002). *RADS-2, Reynolds Adolescent Depression Scale professional manual*. Lutz, FL: Psychological Assessment Resources.

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

Rogers, R. (1984). *Rogers criminal responsibility assessment scales (RCRAS) and tests manual*. Odessa, FL: Psychological Assessment Resources.

Rogers, R. (1992). *Structured Interview of Reported Symptoms*. Odessa, FL: Psychological Assessment Resources.

Rogers, R. (1997a). Children and deception. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 153-166). New York: Guilford Press.

Rogers, R. (1997b). Introduction. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 1-19). New York: Guilford Press.

Rogers, R. (1997c). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 398-426). New York: Guilford Press.

Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception,* (3rd ed.). New York: Guilford Press.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Reported Symptoms professional manual*. Odessa, FL: Psychological Assessment Resources.

Rogers, R., Hinds, J. D., & Sewell, K. W. (1996). Feigning psychopathology among adolescent offenders: Validation of the SIRS, MMPI-A, and SIMS. *Journal of Personality Assessment*, *67*, 244-257.

Sattler, J. M. (2002). *Assessment of children: Behavioral and clinical applications* (4th ed.). San Diego: Author.

Smith, G. P. (1992). *Detection of malingering: A validation study of the SLAM test.* Unpublished dissertations. University of Missouri – St. Louis.

Stein, L. A. R., & Graham, J. R. (1999). Detecting fake-good MMPI-A profiles in a correctional facility. *Psychological Assessment*, *11*, 386-395.

Stein, L. A. R., & Graham, J. R. (2001). Use of the MMPI-A to detect substance abuse in a juvenile correctional setting. *Journal of Personality Assessment*, *77*, 508-523.

Stein, L. A. R., & Graham, J. R. (2005). Ability of substance abusers to escape detection on the Minnesota Multiphasic Personality Inventory-Adolescent (MMPI-A) in a juvenile correctional facility. *Assessment*, *12*, 28-39.

Stein, L. A. R., Graham, J. R., & Williams, C. L. (1995). Detecting fake-bad MMPI-A

    profiles. *Journal of Personality Assessment*, *65*, 415-427.

Stouthamer-Loeber, M. (1986). Lying as a problem behavior in children. A review.

    Clinical Psychology Review, *6*, 267-289.

Stouthamer-Loeber, M., & Loeber, R. (1986). Boys who lie. *Journal of Abnormal Child*

    Psychology, *14*, 551-564.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston,

    MA: Pearson.

United States Public Health Service. (2000). *Report of the surgeon general's conference*

    on children's mental health: A national action agenda. Washington, DC:

    Department of Health and Human Services.

Wiener, D. N. (1948). Subtle and obvious keys for the MMPI. *Journal of Consulting*

    Psychology, *12*, 164-170.

Wimmer, H., Gruber, S., & Perner, J. (1984). Young children's conception of lying:

    Lexical realism-moral subjectivism. *Journal of Experimental Child Psychology*,

    37, 1-30.

Wrobel, T. A., Lachar, D., Wrobel, N. H., Morgan., S. T., Gruber, C. P., & Neher, J. A.

    (1999). Performance of the Personality Inventory for Youth validity scales.

    Assessment, *6*, 367-376.

Appendix A

Integrity Questionnaire

Please answer the following questions about the work you just finished. There are no right or wrong answers. Just mark the answer that best describes what you think. *Please circle your answer.*

1. Did you understand the instructions?   **YES**   **NO**

2. When you answered the questions, were you thinking about how the person described in the instructions would think or feel or behave?
                                                                                                YES   **NO**

3. When you answered the questions, were you thinking about how the person described in the instructions would answer the questions?
                                                                                                YES   **NO**

4. When you answered the questions, were you thinking about how you think or feel or behave?   **YES**   **NO**

5. When you answered the questions, were you thinking about your own answers to the questions?   **YES**   **NO**

6. Do you think that you answered the questions like the person described in the instructions would probably answer the questions?
                                                                                                YES   **NO**

6. Do you think your answers describe how you think, feel, or behave?
                                                                                                YES   **NO**

7. Do you think your answers describe how a pretend person thinks, feels, or behaves?
                                                                                                YES   **NO**