11-10-2017

Polynomial Regression Analyses of Parental Desirable Responding

Megan Elizabeth Lilly
Louisiana State University and Agricultural and Mechanical College, firstone87@gmail.com

POLYNOMIAL REGRESSION ANALYSES OF PARENTAL DESIRABLE RESPONDING

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Megan Elizabeth Lilly
B.S., University of New Orleans, 2009
M.A., Louisiana State University, 2015
December 2017

Table of Contents

Abstract

Child assessments have consistently yielded poor agreement between parent-child self-report measures and considerable research has failed to isolate the underlying source of the discrepancies. Validity testing has successfully been used to differentiate credible from non-credible performance, recognize feigned symptoms, and detect underestimates of ability in adults for decades (Sherman, 2015). Despite these potential benefits, attention to the importance of validity testing in the child assessment literature has been largely neglected. Many psychologists believe children are not capable of feigning or exaggerating their symptoms and lack the underlying motivations and incentives for deception found in adult assessments. Whereas, more research is needed to determine the legitimacy of these beliefs, the reliance on parental reports when assessing child symptoms provides an opportunity to integrate symptom validity testing into child assessment practices by using a population where it has already been proven effective. The current study utilized polynomial regression analyses in order to determine whether an embedded symptom validity measure could predict symptom agreement across informants. Participants included 220 children (42.7% female) ages 11-17 ($M = 12.52$, $SD = 1.43$) and their parents. Four separate polynomial regression analyses were conducted across symptom constructs for which children are commonly referred for outpatient treatment services (depression, anxiety, inattention, and aggression). Results showed that parental desirable responding was detected across several constructs and predicted informant discrepancies in anxiety and aggression symptoms.

Introduction

Despite the recent trend towards adopting a more medical or biological model for explaining mental illness (Cloniger, 1999; Widiger & Clark, 2000), the field of psychology intrinsically lacks the absolute true scores (i.e., scores which exist independent of the measurement process used; DNA, pregnancy tests) that exist throughout the field of medicine. Instead, psychological assessment is concerned with the measurement of underlying constructs and overt or covert characteristics that represent some quantity of the desired attribute (Cantwell, 1996). The possibility for measurement error therefore exists, whereby an individual may provide less than accurate responses about themselves and distort understanding of the underlying construct.

Since the first attempts to standardize and unify assessment methods in the diagnosis of psychological disorders, child assessment procedures comparatively have fallen behind adult practices. The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; American Psychiatric Association), published in 1952, represented the first formalized and codified nomenclature for psychological disorders and included 107 separate diagnoses for adults but only one diagnosis (i.e., adjustment reaction of childhood/adolescence) that could be applied to children. These early diagnostic evaluations rarely included the child themselves but relied solely on information provided by parents (Herjanic, Herjanic, Brown, & Wheatt, 1975; Loeber, Green, & Lahey, 1990; Williams, McGee, Anderson, & Silva, 1989). The first study demonstrating that children could serve as valid and reliable sources of their own feelings and behaviors during a psychological assessment interview was not conducted until 1968 (Rutter & Graham).

The methodology for conducting psychological assessments with children has changed significantly in the last 60 years. Modern typical practice includes a multi-method multi-informant approach that integrates information from a variety of sources familiar with the child's psychosocial functioning. Whereas the practice of gathering information from multiple sources has the potential to provide a broader picture of the psychosocial sequelae contributing to the child's current functioning, it also introduces additional opportunities for error from each source. Therefore, each new informant represents a unique source of error variance that must be accounted for in the child psychological assessment process.

The proceeding literature review introduces the various sources of error and biases that have been documented and the field's reliance on the downward projection of theoretical models used with adults to develop child assessment procedures. The review begins with a discussion of the consistently low rates of agreement between parent-child informants during child assessment practices and highlights the factors that have been associated with influencing concordance. Next, it reviews the relevant clinical, research, and statistical practices that are used to integrate information from multiple sources. Subsequently, the historical importance of validity testing in accurately assessing psychological constructs with adults is provided. In addition, information is presented with examples from forensic evaluations regarding the role of parental impression management in the child assessment process. Next, attribution theory and issues associated with the stigma of mental illness are introduced as factors that can influence parental responses during assessments of their children. Finally, specific applications for validity testing are proposed and the rationale for the current study is provided.

Multiple Informant Agreement

 The assessment of childhood psychopathology represents a unique process in which, for a majority of the population, initiation of services is not sought by the individual in need of help (Kazdin, 1989; Stanger & Lewis, 1993; Yeh & Weisz, 2001). Instead, most assessments originate through the process of parental referral for identified concerns. In addition, historical practices relied almost exclusively on parental reports for child assessments and the gradual inclusion of child self-reports often revealed low correspondence between informants (Lapouse & Monk, 1958). In a seminal meta-analysis investigating agreement between parent-child informants, Achenbach, McConaughy, and Howell (1987) found an overall correlation of $r = .25$ across studies. Beyond parent-child agreement, the meta-analysis included 119 studies spanning from 1960 to 1985 that reported correlations between multiple combinations of informants regarding the child's behavioral and emotional difficulties (e.g., parents, teachers, peers, clinicians, child; Achenbach et al., 1987). It is important to highlight that the methodology for child assessments varied dramatically across the time period for this review. Specifically, the first study to demonstrate that children were able to reliably provide information about themselves during a psychological interview, was not published until eight years into the date range for the meta-analysis (Rutter & Graham, 1968). Furthermore, early assessments exclusively utilized an interview format for child informants and their inclusion developed slowly, representing only 33 of the 119 publications (Achenbach et al., 1987).

 In response to the low levels of agreement between parents and their children on assessment measures, Yeh and Weisz (2001) evaluated 381 parent-child dyads to determine their level of correspondence solely on the identified referral problem. Children and their caregivers were interviewed separately and respectively asked to report "the major problems for which you

3

feel you need help/your child needs help" (Yeh & Weisz, 2001). Results indicated that of the 381 parent-child pairs, 177 (63%) failed in their agreement on a single presenting problem (Yeh & Weisz, 2001). Recently, an updated review of children's cross-informant correspondence was conducted spanning the 25 year interim (from 1989 to 2014) since the meta-analysis conducted by Achenbach and colleagues (1987) that found an overall correlation of $r = .28$ across 341 studies (De Los Reyes et al., 2015). Notably, the authors attributed the significant increase in number of studies both to greater acceptance of the child's role in the assessment process and to the development of a large body of self-report rating scales that were not available during the earlier meta-analysis (De Los Reyes et al., 2015). Despite these changes, results across all three large-scale reviews reveal low-to-moderate correspondence between informants has remained consistent over time. The problem of low informant correspondence has received significant attention throughout the literature; however, to date no overarching theory has emerged that adequately or consistently explains the discrepancies. Current measurement of psychological constructs relies on integrating a number of person-centered variables across situational contexts, which can often lead to variation in accurate measurement. Whereas the inclusion of multiple informants for child assessment practices may allow for the collection of additional data, it may also further complicate the already intricate process of diagnosis. Assessing the specific characteristics of a particular child may vary not only based on child-centered variables, but also on the added characteristics of a particular informant. In addition, as no standard exists for the measurement of any specific psychological construct, assessment practices can also vary based on the measure chosen for use, as well as situation and setting specific determinants. Therefore, research investigating these various sources that may contribute to the discrepancies across informants is required in order to facilitate accurate interpretation of the data.

Agreement Based on Parent Factors

As parents are the most commonly utilized additional informant in child assessment practices, investigations into parent-centered variables and their contribution to informant discrepancies has received significant investigative attention. Parent's responsibility for initiating services for their child may provide a unique source for contamination, as their description of the child's difficulties is often used to guide selection and construction of the subsequent assessment procedures. Specifically, based on the parent's identified concerns, clinicians may form initial hypotheses regarding their expectations for the case that can bias their decision making during the evaluation (Croskerry, 2003; Thammasitboon & Cutrer, 2013). Informant discrepancies may be exacerbated when clinician's utilize a flexible battery approach to select construct specific measurement scales rather than maintaining broader methods of differential diagnosis (Garb, 1998; Groth-Marnat, 2009). Therefore, research studies examining parent-specific characteristics that may vary the accuracy of their reports are critically important given the ongoing reliance on their perspectives in making clinical decisions during child assessments.

Previous research examining divergence between parent-child reports has investigated whether variations may be associated with mental health concerns displayed by parents. For instance, one study that utilized the Dutch version of the *Symptom Checklist 90 Items-Revised* (SCL-90-R; Derogatis, 1992) found maternal psychopathology accounted for up to 20% of the variance across raters (Kroes, Veerman, & De Bruyn, 2003). A number of studies have analyzed the influence of specific parental symptoms and disorders on informant discrepancies. For example, depressed mothers have consistently been found to over-report their children's emotional problems (e.g., Boyle & Pickles, 1997; Collishaw, Goodman, Ford, Rabe-Hesketh, & Pickles, 2009; Durbin & Wilson, 2012; Fitzmaurice, Laird, Zahner, & Daskalakis, 1995; Offord

et al., 1996). The consistency of this finding led to the development of the *depression-distortion hypothesis*, the tendency for depressed adults to display biases that may cause them to over endorse emotional and behavioral difficulties in their children (Matthews & MacLeod, 2005; Müller & Furniss, 2013; Richters, 1992). Furthermore, one study found even higher maternal reports of their child's emotional difficulties when the mother exhibited current depressive symptoms as compared to mothers without current symptoms but the same number of previous depressive episodes (Najman et al., 2000). In addition, high levels of maternal anxiety (Briggs-Gowan, Carter, & Schwab-Stone, 1996; Durbin & Wilson, 2012), maternal alcoholism (Foley, Rutter, Angold, & Pickles, 2005), and both maternal and paternal stress have been positively correlated with an increase in reported levels of both internalizing and externalizing symptoms in their child (Grills & Ollendick, 2003; Jensen, Xenakis, Davis, & Degroot, 1988; Kolko & Kazdin, 1993; Renk, Roddenberry, Oliveros, & Sieger, 2007). Finally, studies have generally shown that socioeconomic status does not significantly influence informant correspondence (Chi & Hinshaw, 2002; Kolko & Kazdin, 1993; Salbach-Andrae, Klinkowski, Lenz, & Lehmkuhl, 2009; Treutler & Epkins, 2003) unless also moderated by an increase in levels of parental stress associated with their lower income (Duhig, Renk, Epstein, & Phares, 2000).

Agreement Based on Child Factors

There has been a dramatic increase in the child's role within the assessment process over the last few decades. Studies investigating child-specific characteristics that may influence agreement have yielded the same poor results found across parent-specific characteristics. For example, researchers have found no relationship between child gender and discrepancies (Achenbach et al., 1987; Choudhury, Pimentel, & Kendall, 2003; De Los Reyes et al., 2015; Engel, Rodrigue, & Geffken, 1994; Hodges, McKnew, Cytryn, Stern, & Kline, 1982).

Furthermore, studies examining whether the age of the child influences agreement have not established a consistent pattern. Specifically, researchers have found that younger children both exhibit lower levels of agreement with their parents (Breton et al., 1995; De la Osa, Ezpeleta, Domenech, Navarro, & Losilla, 1997; Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985; Rapee, Barrett, Dadds, & Evans, 1994; Schwab-Stone, Fallon, Briggs, & Crowther, 1994; Silverman & Eisen, 1992) and higher levels of agreement with their parents (Thompson et al., 1993). Early studies suggested that younger children were unable to fully understand the questions or provide insightful responses about their own feelings and behaviors (Breton et al., 1995; Edelbrock et al., 1985; Rapee et al., 1994; Schwab-Stone et al., 1994). However, more recent conceptualizations where age of the child was grouped by developmental variables such as reading ability or grade level rather than the child's date of birth have found no significant discrepancies (Carlston & Ogles, 2006; Yeh & Weisz, 2001).

Among older children, parent-adolescent agreement rates are frequently influenced by the type of symptom being measured and the developmental shift towards independence that results in less frequent contact with parents, and increasing reliance on peer groups rather than the family unit for emotional support (Brown, 1990; Edelbrock et al., 1985; Sourander, Helstela, & Helenius, 1999). For example, researchers have suggested that adolescents are more reliable reporters than their parents for measuring internalizing symptoms (Achenbach et al., 1987; Andrews, Garrison, Jackson, Addy, & McKeown, 1993; Cantwell, Lewinsohn, Rohde, & Seeley, 1997; De Los Reyes et al., 2015); however, they frequently endorse more depressive symptoms, which results in greater parent-adolescent disagreement (Handwerk, Larzelere, Soper, & Friman, 1999; Kolko & Kazdin, 1993; Renouf & Kovacs, 1994; Seiffge-Krenke & Kollmar, 1998; Tarullo, Richardson, Radke-Yarrow, & Martinez, 1995; Verhulst & Van der Ende, 1992).

7

Other Sources of Error

      Agreement based on type of assessment. Overall, despite decades of studies investigating parent-child informant discrepancies, no single parent or child specific factor has been identified that can adequately explain the incongruence. Given the consistency of discrepancies across these individual levels, it is likely that a more nuanced explanation is required by identifying factors that may influence both informants. Self- and parent-report rating scales have consistently yielded poor reliability between raters (Achenbach et al., 1987; De Los Reyes et al., 2015). Initially this was thought to be a product of test forms that were too dissimilar across raters; however, the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach et al., 1987; Achenbach & Rescorla, 2001), which includes parallel test forms also consistently yields poor agreement correlations between parents and children (e.g., $r =$ .25; Achenbach et al., 1987; Achenbach & Rescorla, 2001). Studies investigating agreement on structured and semi-structured diagnostic interviews have yielded higher levels of inter-rater and test-retest reliability than self-report data; however, agreement between informants remains low (Achenbach et al., 1987; De Los Reyes et al., 2015; Grills & Ollendick, 2002; Jensen et al., 1999). Using the interpretations of the Kappa statistic (i.e., κ) proposed by Hodges and Cools, values of κ < .50 are poor, κ = .50-.70 are acceptable, and κ >.70 are good or excellent (1990). A meta-analytic study of child interview data found parent-child informant agreement rates ranged from κ= .15- .75 across multiple assessments (Sylvester, Hyde, & Reichler, 1987).

      Agreement based on symptom severity. Some researchers have shown that increases in the severity of children's problem behaviors correlate with a child-reported decrease in severity and a concurrent increase in severity ratings reported by their parents (Handwerk et al., 1999; Herjanic & Reich, 1982). An explanation for this pattern has been suggested where children may

under report behavioral difficulties as a means of escaping possible punishment from their

parents (Jensen et al., 1999). Further studies have shown that when assessing for anxiety

symptoms in their children, parental agreement increases in more severe cases as opposed to

cases with mild to moderate symptom severity (Canino et al., 1987). One rationale proposes that

increases in symptom severity related to internalizing symptoms allows for more observable

behavioral difficulties for parents to report (Angold et al., 1987). In contrast, others have found

no difference in parent-child agreement based on internalizing symptom severity (Klein, 1991).

Combining the Data

In addition to the complexity involved in understanding and interpreting the different

sources of information in child assessment, there is also strong variability applied to the practice

of integrating the data across informants to produce a final diagnosis. Table 1 introduces the

most commonly cited integration practices for psychologists in the literature. Table 2 provides

specific theoretical models used to account for these practices across both research and clinical

settings. For example, according to the criteria set forth in the current iteration of the *Diagnostic

and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; APA, 2013) in order for a child to

receive a diagnosis of Attention-Deficit/Hyperactivity Disorder (ADHD) they must meet a

minimum symptom count of six in either the hyperactive/impulsive or inattentive domains.

Table 1.
Common Methods for Symptom Diagnosis during Child Assessments

Method	Description	Author(s)	Potential Problem
Lenient Criteria ("Or")	Symptoms are valid regardless of informant	Bird et al. (1992) Piacentini et al. (1992)	May lead to inflation in prevalence rates across disorders
Strict Criteria ("And")	Symptoms must be endorsed by multiple informants to be valid	Wolraich et al. (2004)	May not account for differences in opportunity to observe certain behaviors across settings

By utilizing multiple informants during a child assessment, a clinician might learn, through parent-report that the child has difficulties listening when spoken to directly, difficulty organizing materials, problems with making careless errors on schoolwork, and trouble related to losing materials for class.

Table 2.
Models for Integrating Discrepant Reports across Clinical and Research Settings

Model Name	Source	Description
Accuracy Model	Muller et al. (2014) Richters (1992)	All informants ratings contribute accurately to measure the child's psychopathology
Distortion Model	Muller et al. (2014) Richters (1992)	Assumes bias in a parent's ratings of their child's psychopathology based on existing parental psychopathology
Combinatory Model	Fergusson et al. (1993)	Parental psychopathology has a causal effect on increases in child psychology and parental psychopathology may lead to increased reports of child symptoms based on biases in parental ratings
Optimal Informant	Holmbeck et al. (2002)	Designates a specific informant to represent the ideal source of information for the desired variable
Separate & Simultaneous	Kraemer et al. (2003)	Separately reports all of the data from each informant at each measurement point throughout the study, regardless of levels of agreement
Aggregation	Kraemer et al. (2003)	Adjustment decisions made on the judgment of the individual researcher after viewing the results
Mix and Match	Kraemer et al. (2003)	Mathematical equation combining the separate influences of the observed trait, the context in which the trait is observed, the characteristics of the individual informant that may bias their observations, and random measurement error

In addition, the child may self-report difficulty finishing schoolwork and chores, making careless errors on schoolwork, being easily distracted by things going on around them, and losing materials needed for class. Currently, psychologists lack a consensus regarding a standardized approach for integrating these discrepant reports between parents and their children. Therefore, a child who was assessed by two different psychologists and provided the same information to each may be assigned dissimilar diagnoses based solely on the preferred integration procedure of the assessor. Most notably, these procedures are often based on the individual preferences of a particular clinician, such as giving higher priority to self-reported symptoms when presented with discrepant reports about anxiety symptoms (Ferdinand, van der Ende, & Verhulst, 2004; Klein, Dougherty, & Olino, 2005).

An alternative strategy that was originally developed for making diagnostic judgments across adult assessments is the *best-estimate* procedure. This method outlines a set of predetermined guidelines both for data collection, using multiple sources and informants, and diagnosis so that any disagreements found from the data are subjected to an integration algorithm (Klein, Ouimette, Kelly, Ferro, & Riso, 1994; Spitzer, 1983). It was initially developed to reduce clinician biases when integrating discrepant information, especially when an individual informant was known to be unreliable.

Development of the specific guidelines for the process varies widely across studies; however, most researchers reported that they relied on thorough independent review of multi-informant data and that diagnostic formulations were made based on predefined formulations rather than relying on clinical preferences or judgments (Garb, 1998; Pilkonis, Heape, Ruddy, & Serrao, 1991; Spitzer, 1983). One of the first estimation processes that included the actual set of guidelines in their publication was used for assessing adults with personality disorders (Klein et

al., 1994). The authors specified that initial data collection included both the identified patient and at least one close family member who both provided information using multiple interview and self-report forms, and the identified patient was also asked to provide all available medical and treatment records (Klein et al., 1994). Following data collection, the authors reported using a set of hierarchically weighted guidelines for making determinations if there was disagreement across any source or format of the assessment information (Klein et al., 1994). For example, one of the guidelines specified that a response to a direct question should be weighted lower if discrepant from another informant if they had spontaneously provided their response during an interview (Klein et al., 1994). In this way, each response is analyzed for differences and resolved before making a final diagnosis. The best estimate approach has consistently exhibited high reliability with adult populations (Garb, 1998).

Replication within child assessments has not consistently shown any improvement over existing practices for data integration (Jensen-Doss, Youngstrom, Youngstrom, Feeny, & Findling, 2014). Furthermore, researchers have suggested that by requiring analysis at every finite level of data collection, the approach requires significantly more time and resources compared with less structured methods (Pellegrino, Singh, & Carmanico, 1999). Regardless of the method chosen for integration, without clear theory-driven guidelines and standards, child assessments will continue to lack the reliability of their adult assessment equivalents. As epidemiological data and the establishment of evidence-based treatments rely on the assumption of proper classification, it is clear that concurrent evidence-based assessment procedures are needed.

Measurement of Informant Discrepancies

   Difference scores. The particular statistical methodology selected for operationalizing and testing congruence has also been associated with increases is parent-child disagreement across studies. Informant discrepancies have historically been measured using a variety of difference score techniques including raw difference, absolute difference, and standardized difference scores (De Los Reyes & Kazdin, 2004; Laird & Weems, 2011). Perhaps the simplest technique used for evaluating informant discrepancies is to calculate a *raw difference score*, also referred to as a *directional difference score*. The score is calculated by simply totaling the raw scores provided by each informant and then subtracting one informant's total score from the other informant's total score. Raw difference scores have been used in past studies to demonstrate the superiority of one informant over another to investigate hypothesized sources of informant discrepancies (Achenbach et al., 1987; De Los Reyes et al., 2015). *Absolute difference scores* are created by first calculating the directional difference score and then taking the absolute value of the result, thereby eliminating the directional (i.e., positive or negative) sign. In contrast to raw difference scores, absolute difference scores have been used by researchers as a measure of the overall level of agreement between informants, rather than as a measure of differentiating which informant is the better reporter of a particular construct (De Los Reyes et al., 2015). *Standardized difference scores* were popularized following the work of De Los Reyes and Kazdin and have represented the norm for analyzing informant discrepancies for the past decade (2004). Similar to raw and absolute difference scores, parent-child reports are constrained so that they represent equivalence (i.e., parent-child reports should theoretically be exactly the same) and results from applying the standardized difference scores are examined as evidence of existing informant discrepancy (De Los Reyes & Kazdin, 2004; Griffin, Murray, & Gonzalez,

1999). In contrast to raw and absolute difference score calculations, standardized difference scores are first converted to z-scores from their raw score total before subtracting one informant's score from the other informant's score.

While difference scores have been widely utilized for analyzing informant discrepancies, their use has been criticized for a number of reasons but most critically for the inability to describe different patterns of discrepancies, only that one exists (Edwards, 2001; Laird & De Los Reyes, 2013). Specifically, continuing with the parent-child report example, multiple configurations may be present (e.g., parent reporting higher levels of a construct but only at extreme ends, children reporting lower levels of a construct but only until a certain threshold is reached); however, difference scores are only able to detect variation in the overall total score. Therefore, the patterns described may produce the same result as a parent-child dyad where the parent endorsed slightly higher levels of a construct uniformly. Difference scores are constrained to assume a perfect one-to-one linear association exists between parent-child informants that does not allow for multiple patterns to be detected (Achenbach et al., 1987; De Los Reyes & Kazdin, 2015). As research has shown that inequality not agreement between raters generally represents the norm, the results of difference score analyses as measures of informant discrepancies are particularly problematic as conclusions drawn from their results may not accurately encapsulate underlying variations.

Polynomial regression analyses. One statistical technique that has received considerable recent attention in the literature across a number of social science fields is polynomial regression. Specifically, polynomial regression is a type of multiple regression analysis where nonlinear, curvilinear, data can be modeled. Polynomial regression allows for the inclusion of higher order terms, predictor variables raised to an Nth power, that test for non-linear functional relationships

where combined effects and interactions may be examined. Analyses incorporate additional predictors (e.g., quadratic, cubic) into the regression model that allows effects to be estimated beyond the constrained assumption that correspondence between variables must represent a one-to-one linear relationship (Dawson, 2014; Hayes, 2013). Therefore, instead of assuming the same incremental difference, either increasing or decreasing, across the entire measurement range, including a quadratic (i.e., squared) term into the regression model can capture a significant change in the function of the relationship.

Perhaps the most well-known example of this type of relationship in the field of psychology can be found in the Yerkes-Dodson law, whereby increases in performance are associated with increases in arousal but only to a point after which any further increases in arousal are associated with decreases in performance (Yerkes & Dodson, 1908). Furthermore, an equation model that results in a curve rather than a straight line creates additional areas for interpreting the relationship (e.g., steepness of the curve, placement of the peak of the curve, shape of the curve). Specifically, steepness is determined by the absolute value of the quadratic term (i.e., larger values reflect sharper bend) and the shape of a curve may represent either a *convex*, upward slope, or a *concave*, downward slope, effect. A convex relationship occurs when the coefficient for the quadratic term is positive and a concave relationship occurs when the coefficient for the quadratic term is negative (Hayes, 2013).

Social science research is frequently concerned with evaluating the varying effects of agreement and congruence across a number of constructs. Therefore, the increased flexibility offered by polynomial regression analyses has led to a significant increase in its implementation across behavioral disciplines. When applied to the assessment of informant discrepancies, polynomial regression models may predict both the direction and degree of incongruence

between raters allowing for more accurate interpretations compared to difference score approaches. Laird and De Los Reyes first demonstrated the application of these analyses using parent-adolescent informant discrepancies to predict adolescent depression and antisocial behavior (2013). Specifically, results indicated higher levels of antisocial behaviors when both mothers and their adolescents agreed on the presence of their symptoms and lower levels of depression symptoms when they both reported positive parental acceptance (Laird & De Los Reyes, 2013). In addition, the study found higher rates of depression when there was discrepancy between reports of rule-breaking behavior with parents reporting low rule-breaking and adolescents reporting higher frequencies (Laird & De Los Reyes, 2013). These interactions allow for broader conceptualizations about the associations between variables. For example, future predictions might test a theory that parental indifference or detachment may be guiding the association so that the adolescent breaks rules in a bid for parental attention and when the attempts are continuously met with disinterest, depressive symptoms develop. Alternative predictions could also be generated and systematically tested.

Rather than examining the levels of discrepancy as a predictor for some identified outcome, Laird and LaFleur rearranged the equation model so that informant discrepancy could to be treated as the dependent variable and the strength of the agreement or discrepancy between informants would function as the theorized moderator (2016). The study utilized the same database as Laird and De Los Reyes (2013) and examined maternal monitoring, solicitation, and control as predictors of increased mother-adolescent agreement on reports of the adolescent's rule-breaking behavior (Laird & LaFleur, 2016). Results indicated interaction effects where higher maternal reports of monitoring and control through rules predicted higher informant congruence regarding adolescent rule-breaking behavior (Laird & LaFleur, 2016). Also, the

16

congruence between mother-adolescent ratings of adolescent rule-breaking was weaker at lower levels of maternal monitoring and control through rules (Laird & LaFleur, 2016). While these techniques represent an exciting improvement over difference scores for understanding predictors of informant discrepancies, the novelty of their use in the field of applied psychology has led to procedural differences in their application.

Specifically, Laird and De Los Reyes (2013) first applied polynomial regression to informant discrepancies based on an article published by Edwards in 1994. In the original article, Edwards suggested testing coefficient patterns more complex than hypothesized to ensure the models adequately captured any relationships between predictors (1994). Laird and De Los Reyes included cubic coefficients in their model, however, the cubic effects were not described in their results (2013). Subsequently, some researchers have also included the cubic coefficient based on these recommendations but consistently have reported that the cubic terms did not contribute significantly to the models and were subsequently removed (Lai, Beaulieu, Ogokeh, Self-Brown, & Kelley, 2015; Leung, Shek, & Li, 2016; Kelley et al., 2017).

In contrast, other studies have simply chosen not to include cubic terms in their models at all (Human, Dirks, DeLongis, & Chen, 2016; Nelemans et al., 2016; Yaban, Sayil, & Tepe, 2014). These studies are supported by literature that has shown that the current methods for measuring psychological constructs rarely, if ever, exceed quadratic complexity (Dawson, 2014). Thus far, this review has focused on the plethora of research underlying the inherent difficulties with integrating multiple perspectives in classifying latent constructs. Various sources of error variance have been found that may contribute to increased discrepancies across informants and current methods differ on how to accurately detect these differences.

One important obstacle remains neglected. What happens if an individual responds in a manner contrary to the truth? What follows will be a review of this idea, which serves as the central tenet of this study.

Validity Testing in the Child Literature

Validity has always been an integral component of test creation; whether or not a test measures what it is supposed to measure is fundamental to the legitimacy of the practice of psychological assessment. In the adult psychological literature over the last few decades, more than 20 books and over 1000 peer-reviewed articles have appeared concerning the topic (Carone & Bush, 2013; Sweet & Guidotti Breting, 2013). In addition, performance validity tests are generally considered common practice in adult testing in order to determine the veracity of individual responses. Validity testing is used to differentiate credible from non-credible performance, recognize feigned symptoms, and detect underestimates of ability (Sherman, 2015).

In contrast, attention to the importance of validity testing in the child assessment literature has been largely neglected. The first book on the topic of adolescent malingering was not published until the late 1990's (McCann, 1998) and the first neuropsychological review on the topic did not appear until 2004 (Rohling, 2004). Recently, two comprehensive reviews have shown that the majority of psychologists do not view tests of validity as a necessary component of child psychological assessments, despite the abundance of research espousing its importance in the adult literature (DeRight & Carone, 2013; Kirkwood, 2012). In part, this may be due to endurance of the historical belief that children were not capable of feigning or exaggerating their symptoms during psychological assessments (Peterson & Peterson, 2015). Similar beliefs were used initially to exclude children from the entire assessment process by denying that they were capable of accurately describing themselves at all (Rutter & Graham, 1968). These same beliefs

18

were subsequently applied to the reasoning against a need for self-report rating scales, as children would not be able to understand questions about their feelings (Edelbrock et al., 1985). It is essential for researchers and clinicians to identify ways to overcome the endurance of these views. In order to better understand the role that validity testing may play in the future of child assessment practices, it is first important to understand the history of its development in the field of psychology as a whole.

Impression Management, Defensive Responding, Self-Deception

History and development. In 2001, Meyer and colleagues published a comprehensive review of psychological assessment measures wherein they claimed, "we have demonstrated that the validity of psychological tests is comparable to the validity of medical tests and so differential limits on reimbursement for psychological and medical tests cannot be justified on the basis of the empirical evidence" (p.155). Whereas psychologists have recently embraced a more medical model for understanding psychological disorders, diagnostic constructs continue to represent abstract and latent variables that are often not directly observable.

The first widely utilized criterion-referenced measure developed for use in psychiatric assessments was published in 1943 by the University of Minnesota Press (Hathaway & McKinley, 1940, 1942, 1943a). The *Minnesota Multiphasic Personality Inventory* (MMPI; Hathaway & McKinley, 1943b) included 566 true-false items that produced ten clinical scales based on common diagnostic groups from the 1930's. Soon after its publication, Meehl and Hathaway published the first review article discussing the MMPI's inability to accurately be used for diagnosis based on the measure's "susceptibility to faking or lying in one way or another" (1946, p. 525). In addition, they noted that deception had an invalidating influence on personality inventories and formally condemned the *assumption of frankness*, which they

believed to be pervasive throughout the psychological testing community at the time (Meehl & Hathaway, 1946). In the article, the authors bluntly declared, "It almost seems as though we inventory-makers were afraid to say too much about the problem because we had no effective solution for it, but it was too obvious a fact to be ignored so it was met by a polite nod" (Meehl & Hathaway, 1946). Therefore, it is clear that while early researchers may have understood the distorting influence that could occur when individuals failed to respond accurately, they lacked evidence-based mechanisms for dealing with the problem.

What followed were a series of investigations that ultimately led to modern standards for validity testing in the field of psychology. Some of the first studies attempting to understand the problem were designed to test whether individuals were even capable of faking various characteristics when asked (Bernreuter, 1940; Olson, 1936). One study administered a personality inventory to separate groups of teachers, one group was prompted to answer the items as though they would be given to a potential employer during an interview and the other was provided neutral instructions to answer honestly (Olson, 1936). Results indicated that teachers in the job group earned significantly higher scores across items related to extroversion and self-sufficiency compared to the teachers in the neutral group (Olson, 1936). Bernreuter (1940) designed a comparable study using an undergraduate college student sample and found similar results where students in the test group were able to present themselves as more dominant and self-sufficient than their counterparts in the control group.

Multiple early researchers found that anonymity during testing could consistently produce inflated responses by having respondents either sign their name or leave their names off of testing materials (Maller, 1930; Spencer, 1938). Whereas these studies did not provide solutions for dealing with falsified test data, they were responsible for creating the necessary

experimental evidence proving that individuals were capable of distorting their response patterns. Early correction methods for informant distortions were highly variable. One researcher suggested that items on personality measures should be written in a manner that less obviously revealed the significance of responding in a certain way (Meehl, 1945). This technique closely resembles the modern idea of *face validity* in that a particular test may more or less appear to measure what it actually measures.

Attempts to disguise the significance of items on personality measures met with initial resistance, as items were chosen for their obvious practicality in measuring a particular trait and integrating subtle items was thought to require unreasonably long tests (Meehl & Hathaway, 1946). One strategy suggested as a more useful mechanism for disguising items was to reverse the wording across approximately half of the test so that an affirmation of any particular response would not necessarily be considered negative for all the items (Meehl, 1945). Maller utilized his prior research investigating anonymity to correct for deception by reformulating his test questions in the third person, thus allowing the individual to indicate sameness or difference with a description of characteristics (Maller, 1932). One method for dealing with informant distortions stemmed from the concept that individuals may be responding in a socially desirable way because they felt that by indicating their responses in writing, they were creating a permanent account of their faults (Maller, 1932). Instead, the investigators allowed the respondent to indicate their answers by displaying cards that had the individual test items printed on them (Maller, 1932). The popularity of this technique led to its inclusion as an administrative option on the MMPI (Hathaway & McKinley, 1943b).

Following these early attempts at eliminating the potential for personal biases in self-report ratings, subsequent trends were directed more towards identifying and adjusting for the errors after they were determined to have occurred. One researcher summarized the shifting attitude when he said, "Might it not be more effective to recognize at the outset that such tests have certain limitations that can never be completely circumvented and then go on to the measurement of these limiting factors themselves…?" (Rosenzweig, 1934, p. 400). He further suggested that self-report measures could be improved by including a set of items measuring an individual's idealized version of themselves, thereby creating a basis for future comparison should corrections be warranted (Rosenzweig, 1934).

An early attempt at creating embedded validity items was created utilizing a subset of questions designed to produce an index of the individual's *readiness to confide* (Maller, 1932). Follow-up studies found the index did not significantly improve the measure's validity and it was removed from subsequent versions (Meehl & Hathaway, 1946). Instead, in a revised version of his measure Maller utilized what he termed an *inconsistency score,* where each test item was repeated a second time but phrased in reverse so that significant discrepancies between item pairs could be identified and analyzed (1944). This technique was also abandoned for failure to increase validity, which Maller attributed to the obvious shift in measurement across the test, since each question was negative in the second half and positive in the first (1944). In developing their *Character Education Inquiry*, Hartshorne and May included a number of items that were highly idealized (e.g., I sometimes put off until tomorrow what I ought to do today) and statistically improbable for any individual to consistently fail to endorse (1928). Under endorsement of a significant number of these items, responding *Not True*, was considered as valid evidence that the individual was engaging in deception in a way that overemphasized their

positive attributes, defined as *faking good* (Hartshorne & May, 1928). In addition, this method had the added benefit of being able to detect not only those individuals who wish to make themselves look good, as with the above example, but using the same technique could also detect individuals who were attempting to respond in the opposite direction where they *over endorsed* their negative attributes. Using the same item described earlier in this manner, a respondent who answered *Very True* to this and a significant number of similar items could be identified as *faking bad* (Humm & Wadsworth, 1935). Therefore, an opposing bi-directional model of deception was established, as an individual who either over-or-under endorsed a considerable number of these items, could be identified as either faking good or faking bad. In follow-up analyses, Humm and Wadsworth found approximately 30% of individuals responded to these items in a way that suggested they were either faking good or faking bad (1935). The purported success of this method in detecting an individual's deliberate attempts to respond deceptively led to its adoption and further expansion on subsequent versions of the MMPI (Hathaway & McKinley, 1943b). In addition, it established the two-factor model for measuring desirable responses that is still used today.

      Modern theories and definitions. Since the early attempts at developing procedures for measuring the accuracy of responses on self-report measures, validity testing has remained a central concern and studies have been continuously published endeavoring to refine the methodologies for identifying adult deception (Boone, 2007; Larrabee, 2007; Sweet & Guidotti Breting, 2013). Recently, the American Academy of Clinical Neuropsychology released a consensus statement specifically addressing effort, response bias, and malingering, wherein they concluded that formal assessment of symptom validity was a necessary action for psychologists engaged in clinical practice (Heilbronner et al., 2009). As the methods for measuring invalid

responses have continued to evolve over the last few decades, so too has the terminology for describing these behaviors. Varying terminology is often used interchangeably throughout the literature, which can lead to confusion in clinical practice.

Recent theory integration of socially-desirable responding presented the same two-dimensional structure originally defined by Humm and Wadsworth (1935), wherein on one side an individual may deny undesirable characteristics and on the other they may enhance those characteristics they view as undesirable (Schutz, 1998). In addition, socially desirable responding may be further differentiated based on the audience, whereby *impression management* represents deliberate attempts at altering presentation of the self to others, and *self-deception* represents biased internal self-description (Paulhus, 1984). Humm and Wadsworth (1935) originally used the term faking good to describe an individual's tendency to present themselves in a highly idealized and socially desirable manner, while minimizing responses seen as less than desirable. Today, the term faking good is also referred to as impression management, *positive impression management, defensiveness,* or *defensive responding.* The other term proposed by Humm and Wadsworth, faking bad, describes an individual's tendency to present themselves in an overly negative and socially undesirable manner (1935). Today, the term faking bad is also used as an umbrella term to describe *feigning, exaggeration, fabrication,* and *malingering*. In contrast to the interchangeable vocabulary used to describe faking good responses, the faking bad terms actually represent distinctly different concepts. *Feigning* is purposeful endorsement of a symptom or condition where none are present (Sherman, 2015). *Exaggeration* is purposeful embellishment or aggrandizement of actual symptoms resulting from a real condition (Sherman, 2015). Finally, *fabrication* refers to the purposeful creation of symptoms as evidence for a condition that an individual does not have (Sherman, 2015).

Diagnostic conditions. The *DSM-5* recognizes two conditions related to deceptive behaviors (APA, 2013). The first condition, factitious disorder, is the deceptive falsification of physical or psychological symptoms without evidence for external incentives. Factitious disorder falls within the somatic symptom and related disorders category of diagnoses and includes a sub-diagnostic option, factitious disorder imposed on another, to identify individuals who engage in deceptive falsification in another without evidence for external incentives (APA, 2013). The second condition is not considered a mental disorder in the *DSM-5*; rather, it falls within the category of other conditions that may be a focus of clinical attention. Malingering is the deceptive exaggeration or production of symptoms with evidence for external incentives (APA, 2013). Therefore, the primary difference between the two conditions is the presence or absence of external incentives. A second distinction should be noted, emphasizing how the conditions are resultantly classified within the *DSM-5* (APA, 2013). The lack of an obvious reason for the deception is considered a diagnosable psychological disorder, while an individual engaging in deception because they could perceptibly gain something from doing so, is merely considered noteworthy.

Specifically within child assessment practices, the malingering condition can be applied to youth; however, the examples used to define external incentives (i.e., avoiding military duty, avoiding work, obtaining financial incentives, evading criminal prosecution, obtaining drugs) were clearly developed for an adult population (APA, 2013). In addition, the four conditions that are provided by the *DSM-5* as valid means of identifying malingering are derived from the Slick, Sherman, and Iverson (1999) criteria that were created through research with adults. The authors of these criteria briefly discussed application to youth populations; however, they did not provide insight as to how child clinicians could detect the occurrence of malingering or what to do if

malingering was in fact detected (Slick et al., 1999). Furthermore, as their research did not actually include child data, they were unable to extrapolate whether children were even capable of malingering when there was no obvious opportunity for secondary gain (Slick et al., 1999).

The question of whether children can be officially deemed malingerers in the absence of external incentives was addressed in a survey on the topic distributed exclusively to psychologists engaged in child clinical practice (Rohling, 2004). Only a small number of the surveyed psychologists believed children were capable of engaging in sophisticated deception tactics during assessments and those who did almost unanimously assumed that children lack the incentive to do so (Rohling, 2004). Furthermore, results revealed that a large portion of psychologists view validity tests as unnecessary during child assessments because any attempts by the child to produce less than optimal effort would be inherently naive and easily detected (Rohling, 2004; Walker, 2011). The persistence of these views likely results from failure to apply the concepts within a developmentally appropriate framework. For example, while the external incentives for malingering as defined by the *DSM-5* may not be directly applicable to children, the notion of rewards that are provided for specific behaviors is a concept near universally understood at a young age. In 2010, Kirkwood, Kirk, Blaha, and Wilson conducted a case series involving six children who provided non-credible responses during their assessments and found substantial evidence of developmentally appropriate external incentives (i.e., allowed to stay home from school and play video games, delaying their parents' separation, getting out of having to play sports). When compared to the incentives described in the *DSM-5*, the potential gain for children reflects a clear parallel to those for adults (e.g., avoiding work likened to avoiding school, obtaining financial incentives likened to receiving and playing video games, evading military duty likened to avoiding undesirable responsibilities). When viewed in this

more developmentally appropriate framework, clinicians would likely find it much more difficult to maintain a belief that children are incapable of engaging in deception tactics in order to get out of having to attend school.

Additionally, studies using college students have identified a host of external incentives associated specifically with an ADHD diagnosis (e.g., extra time on tests, assistance with schoolwork and note taking, stimulant medication) that younger children would likely find similarly motivating (Rabiner, 2013). Estimates of malingered ADHD among older teens and young adults range from 10 to 50% (Harrison, Rosenblum, & Currie, 2010; Marshall et al., 2010; Sullivan, May, & Galbally, 2007). Yet, a thorough literature review conducted in 2013 found no existing studies examining children under age 16 and malingered ADHD (Rabiner, 2013).

Desirable Responding

While the motivation for over endorsing symptoms may be partially explained by the presence of external rewards, the impetus for under reporting remains much less clear in child assessment practices. Therefore, understanding the motivation for desirable responding in child assessments will likely benefit from initially targeting parents, given that as adults they represent a population with decades of evidence on the topic. One area of child assessments where previous research has demonstrated evidence for parental desirable responding is in forensic evaluations. Forensic psychological evaluations are often considered "high stakes" assessment environments and the ability to differentiate non-credible responders is often vitally important to the future safety and well-being of the child. In addition, given the significant nature and possible consequences associated with these evaluations there is clear rationale for explaining parental attempts at impression management. Therefore, forensic evaluations represent a unique subgroup of child practice that has successfully utilized validity testing with parents when

conducting assessments. Forensic researchers have demonstrated that parental positive impression management contributes to measurement error during child-custody evaluations following a divorce or separation (Otto & Collins, 1995; Otto, Edens, & Barcus, 2000), general disputes over child-custody (Arce, Farina, Seijo, & Novo, 2015), and during decisions regarding termination of parental rights (Carr, Moretti, & Cue, 2005). In addition, positive impression management has been associated with lower rates of endorsement for substance use in mothers, especially when they were involved in an ongoing case with Child Protective Services (Donohue, Holland, Lopez, Urgelles, & Allen, 2014).

Another area where detection of impression management may play a particularly vital role is within parenting capacity assessments, where a parent may be denied access to their child as a direct result of their responses. Specifically, these evaluations are typically requested when there is a question about a parent's ability to meet their child's basic needs or if their child is deemed potentially at risk for physical/sexual abuse or neglect (Budd, 2001; Kuehnle, Coulter, & Firestone, 2000). Results of a systematic study examining the veracity of parental responses across validity indices revealed that 60% of parents undergoing a parenting capacity assessment exhibited significant elevations for positive impression management, while no individual endorsed significant negative impression management (Carr et al., 2005).

Parental positive impression management during forensic evaluations is likely a result of the high-pressure environment and is used to avoid the significant consequences that may occur based on their responses. However, if validity testing is to be applied to parents as raters of their child's symptoms across a variety of different outpatient settings, a deeper understanding is needed regarding the underlying motivation for why a parent may respond defensively. One possible way of investigating potential motivating factors that should be considered is to design

28

and conduct research studies based on the identification of parents not just as informants of their child's behaviors but as adults to whom the plethora of existing literature on validity testing may be reliably applied. Given the scarcity of research investigating factors associated with impression management for parents as informants for their child's symptoms, examination of the existing rationale for defensive responding in adults may represent the best place to begin developing child-assessment specific theories.

Rationale for Desirable Responding

  Stigma. The term *stigma* is derived from an ancient Greek word that described a burned or cut tattoo given to slaves and criminals so as to delineate them from the rest of the population (Alexander & Selesnick, 1966; Cockeram, 1981; Foucault, 1965; Hinshaw, 2007). Societal norms at the time required these marked individuals to be shunned and avoided in public under penalty of law (Hinshaw, 2007). Since its early use, the term has evolved beyond its physical demarcation to represent a more abstract *mark* of disgrace.

  Erving Goffman first purported a sociological definition for stigma as a discrediting attribute or condition (e.g., physical disability, disease) deviating from the norm (1963). Social psychologists refined this definition to incorporate deviant characteristics and attributes, and with the emergence of the social-cognitive approach within the discipline, later distinguished Goffman's original "attribute" as an individual's view of their social identity within a specified context (Crocker, Quinn, & Steele, 1998). The addition of context helped define a framework of terminology that allowed for auxiliary study of stigma within specific domains (e.g., stereotyping, separation, status loss) and with varying degrees of perceived power (e.g., labeling, discrimination) (Link & Phelan, 2001). These conceptualizations were refined by differentiating between *public stigma*, the negative views and treatment by other individuals, and *perceived stigma*, an

individual's negative cognitions about themselves as a result of the public's negative views (Corrigan, 2004). Gregory Zilboorg wrote extensively about the interaction of public perception and the burgeoning field of psychology starting in the 1930's. Throughout his career, Zilboorg explored the effects of medicine and psychology's relationship with some of the darkest aspects of human history (Hinshaw, 2007). In his 1941 review of the history of psychiatry, he wrote "No new step is ever made and no new discovery is ever brought forth without the shadows of the past hovering over it" (Zilboorg & Henry, 1941). Therefore, Zilboorg understood that stigma and the historical human tendency to ostracize, segregate, and abuse individuals who display aberrant behaviors will forever underlie the public's perception of mental health despite even the most rigorous of scientific advancements.

In 1999, at the end of the twentieth century, Otto Wahl reviewed hundreds of public surveys dating back to the early 1950's, wherein he found that negative views of the mentally ill have persisted across time (Wahl, 1999). In response to the pervasively negative perceptions of the public towards mental illness, Wahl created his own survey designed to measure actual past experiences of stigma and discrimination directly from individuals with mental illness. The survey was completed by more than 1300 individuals between 1996 and 1999 from across all fifty states, Ireland, Wales, Australia, and Canada (Wahl, 1999). All items utilized a five-point Likert scale and respondents were given the opportunity to provide additional written responses of specific examples or to be followed up via interview by one of the researchers (Wahl, 1999). Nine of the survey items asked about individual experiences of stigma (e.g., I have been advised to lower my expectations in life, I have been shunned or avoided by others, I have been treated as less competent) and the remaining twelve items involved direct experiences with discrimination (e.g., I have been denied educational opportunities, I have had difficulty renting an apartment, I have

been turned down for a job). Results of the survey showed only 10% of respondents had never been in a situation where others had said unfavorable or offensive things directly about them and their illnesses. In addition, 95% of interviewees reported that their experiences of stigma and discrimination had had a "lasting impact," and one in ten individuals reported that these experiences made pursuing professional help less likely (Wahl, 1999). Thus, Wahl's survey was of critical importance not only in its inclusion of actual experiences of mental health consumers but in highlighting the influence of these experiences on future treatment seeking behaviors.

Stigma in Children. Despite the preponderance of evidence regarding the relationship between stigma and mental illness in adults, the first systematic review of stigma's relationship to mental health difficulties in children was not published until September 2016 (Kaushik, Kostaki, & Kyriakopoulos, 2016). In their review, Kaushik and colleagues found 31 qualitative papers examining children's stigmatizing views towards their peers who carried mental health diagnoses and seven studies examining adult's views towards children (2016). Despite significant differences in methodological approaches across studies, children with mental health disorders experienced stigmatization by peers at a greater rate than that of undiagnosed peers, children with learning disabilities, and children with debilitating medical conditions (Kauskik et al., 2016). Earlier analyses regarding disorder specific stigmatization found that adults tended to view depression in children more negatively than externalizing disorders (Ohan, Visser, Moss, & Allen, 2013). In contrast, children expressed less stigmatization towards their anxious peers as compared to those with depression or ADHD (Bellanca & Pote, 2013; Coleman, Walker, Lee, Friesen, & Squire, 2009; Dixon, Murray, & Daiches, 2013; O'Driscoll, Heary, Hennessy, & McKeague, 2012; Walker, Coleman, Lee, Squire, & Friesen, 2008). A longitudinal study conducted with children established a positive correlation between child development and peer

31

stigma (Weiss, 1994). As children progressed from early childhood to adolescence they increasingly preferred to distance themselves from peers with mental health labels. Furthermore, by the end of middle school, rankings based on desirability of certain behavioral labels found *crazy* had supplanted *convict* as most intolerable (Weiss, 1994). Given the preponderance of research examining the lasting impact a mental illness label may carry for an individual across their lifespan, it is important to examine the impact of the parental role in ascribing such a label to their child during psychological assessments.

Stigma and Parents. In the context of child assessment practices, the youth is rarely responsible for their own treatment seeking behavior. Therefore, it is important to consider the role of parental perceptions of stigma regarding child mental health diagnoses. In 1963, Goffman first introduced the concept of *courtesy stigma* to describe parental experiences of discrimination and rejection by others who see them as responsible for causing their child's mental illness. Studies examining the link between parental stigma and child psychological disorders have suggested that parents may feel shame and embarrassment as a result of their child's mental health diagnosis (Angermeyer, Beck, Dietrich, & Holzinger, 2004; Corrigan & Miller, 2004; Phelan, Bromet, & Link, 1998). Research on parental perceived stigma of children with specific disorders remains scarce; however, one study found that parents whose child was diagnosed with autism experienced significant levels of associated stigmatization (Gray, 1993).

In order to address the significant lack of empirical research regarding parental perceived stigma, Williams and Polaha created a 17-item parent-report measure using a sample of 347 parents (2014). Two distinct factors, public stigma and self-stigma, were identified and found to predict parental willingness in mental health service seeking behaviors (Williams & Polaha, 2014). Furthermore, parents who reported they would not seek services for their child despite

32

identified psychosocial difficulties frequently endorsed feelings of inadequacy as a parent and fears of unfair treatment towards their child as the primary barriers to help-seeking (2014). These findings reflect those found in the adult literature whereby, especially in rural communities, stigma related to mental illness is frequently cited as the top barrier to individuals seeking services (Bray, Enright, & Easling, 2004; Jameson & Blank, 2007; Williams & Polaha, 2014).

Attribution Theory and Parental Impression Management

Research regarding parental perceptions of stigma related to their role in determining their child's diagnosis remains limited. Attribution theory may provide a possible theoretical framework for future studies. A number of well validated instruments have been created to assess the various dimensions of responsibility that parents feel for their children's behavioral difficulties (e.g., the Parental Locus of Control Scale; Campis, Lyman, & Prentice-Dunn, 1986; Parental Style Attribution Questionnaire; Sobol, Ashbourne, Earn, & Cunningham, 1989; Parent Attribution Test; Bugental, 2004; Parent Cognition Scale; Snarr, Slep, & Grande, 2009). Whereas these measures differ significantly on the specific characteristics associated with parental attributions, they all share the general dimensions of causality and responsibility that represent current theoretical models of attribution theory. The dimension of *causal attributions* includes any explanations for the reasons behind why a particular event occurred (Bradbury & Fincham, 1990). In contrast, *responsibility attributions* pertain to a particular individual's belief surrounding their involvement and accountability for the events' occurrence (Bradbury & Fincham, 1990; Weiner, 1995). As it is applied to child assessment literature, the theoretical constructs of parental causality and responsibility suggest that parents may hold beliefs that their child's behavioral difficulties are a direct result of the parent's actions or dysfunctional qualities (Bradbury & Fincham, 1990).

Attribution theory in medicine. Researchers have yet to investigate whether parental

impression management is correlated with responsibility attributions regarding their child's

mental health symptoms. Despite a paucity of research on the topic within the field of child

psychology, there are a number of studies that have applied attribution theory to parental

impression management across a variety of other medical and health-related disciplines. One

such study evaluated a group of mothers who had previously received in vitro fertilization or

utilized embryo donation in order to conceive a child (MacCallum, Golombok, & Brinsden,

2007). The investigators found that mothers endorsed higher levels of defensive responding on

self-report measures and they reported much higher rates of competency in their role as a parent

and in their ability to handle the stressors associated with the role compared to controls

(MacCallum et al., 2007). Follow-up interviews were conducted to assess possible explanations

for these response patterns and mothers generally indicated under endorsing impairment and

problems as a result of feeling less entitled to complain given their difficulty with initial

conception. Furthermore, mothers reported feeling social pressure to never appear ungrateful and

to view stressors associated with their role as a parent as a positive rather than a burden

(MacCallum et al., 2007).

Another study investigating 209 parents and their preschool-aged children, found higher

levels of parental desirable responding were associated with decreased parental reports of the

frequency of unhealthy snack and fast food consumption in their overweight children (Radnitz &

Todd, 2016). In addition, no association was found between parental desirable responding and

parental reports of unhealthy food consumption for children whose weight fell within the average

or better range. Furthermore, while the authors hypothesized a related association would be

found between higher levels of parental desirable responding and increased reports of the

frequency of their overweight children's healthy food (e.g., fruits and vegetables) consumption, these results were not significant (Radnitz & Todd, 2016). Therefore, the researchers postulated that parents may not necessarily feel compelled to prove they exhibit positive parenting practices; rather, parents of overweight children may under report the frequency of unhealthy eating behaviors as a means of minimizing their perceived contribution to and feelings of responsibility for their child's weight problems (Radnitz & Todd, 2016). Previously, a qualitative study was conducted via open-ended interviews with mothers of children identified as obese in order to assess potential factors that may contribute to maternal patterns of desirable responding regarding their child's weight (Jackson, Wilkes, & McDonald, 2007). The findings revealed that the mothers were acutely aware that their children were overweight; however, they expressed feeling the need to minimize their child's weight problem to others as a result of negative societal stigma and experiencing frequent judgment and blame about being a bad mother (Jackson et al., 2007). In addition, the mother's reported high levels of frustration about their perceived lack of ability to intervene effectively and anxiety about the long-term impact of their child's weight on their self-esteem, social and romantic relationships, and emotional development (Jackson et al., 2007). The researchers postulated that maternal defensive responding was likely a product of these mixed emotional experiences as mother's balanced feelings of responsibility to actively seek treatment for the problem while simultaneously fearing that doing so may increase their distress by exposing them to increased societal stigma and blame (Jackson et al., 2007).

Southwell and Fox conducted a study with mothers of children ages 6-13 years who were actively enrolled in treatment for weight management and were initially classified as having a Body Mass Index above the 91st percentile (2011). Specifically, the researchers investigated the

association between desirable responding and overall maternal perceptions of their child's weight and found that mothers significantly under assessed their overweight or obese children's size, as well as the severity of the problem overall (Southwell & Fox, 2011). The authors postulated that mothers may under report their child's weight and associated difficulties for the purpose of protecting themselves both from external experiences of stigma and blame and from internal negative feelings of responsibility for their child's nutritional intake. Furthermore, the authors proposed that for some mothers, their child's weight may represent a direct visual indicator of their inabilities as a parent (Southwell & Fox, 2011).

Notably, each of the reviewed studies regarding overweight or obese children were conducted with mothers who voluntarily enrolled their children in weight management treatment programs. Therefore, the findings exemplify the parallel and seemingly contradictory process within child psychological assessments whereby parents actively seek-out services for an identified problem and then subsequently minimize its presence and impact on their child's functioning. Furthermore, these studies indicated that mothers feel intense societal pressure and blame as a result of their child's weight, as well as increased guilt and shame regarding feelings of responsibility for contributing to the problem. These maternal reports are highly correlated with the attribution theory explanation for explaining defensive responding. Whereas the formal codification of attribution theory is a relatively recent creation, it was not present to provide an explanation for historical beliefs regarding child psychological difficulties. Nonetheless, it may retrospectively be applied to explain early views suggesting that the majority of mental illness in children could be blamed on the child's parent, in particular their mother (Hinshaw, 2007). The third iteration of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-III*; APA; 1980) purported that a number of childhood disorders (e.g., early infantile autism, schizophrenia)

were the direct result of faulty parenting practices. Therefore, from its inception child psychology had as official policy, an explicit attribution bias against the parents of children with mental illness. The field of psychology has gradually altered its perception away from the causal relationship between childhood mental illness and parental responsibility; however, survey studies have consistently exposed the general public's resistance to accept a lack of parental causal responsibility (Read & Harre, 2001). Public perception and stigma may therefore continue to reinforce parental attributions of their child's mental health difficulties, despite transition away from such beliefs. The application of attribution theory to stigma of mental illness may help explain a parents need to protect their children as by doing so, parents protect themselves from being ascribed the cause of their child's dysfunction.

Detecting and Measuring Parental Impression Management

The area of research that began in the 1930's in pursuit of reliable instruments and techniques for detecting and measuring respondent veracity has made significant advances over the ensuing decades. These improvements ultimately led to the development of the MMPI that was subsequently updated to produce the revised *Minnesota Multiphasic Personality Inventory-Second Edition* (MMPI-2; Hathaway, McKinley, & MMPI Restandardization Committee) in 1989. This second iteration is generally considered to be one of the most widely used instruments in psychology. Furthermore, its use as a measure for detecting response distortions with parents has been successfully demonstrated in child forensic evaluations previously described in the current study. One review found that the MMPI-2 was used in 90% of all child-custody evaluations over a period of 20-years (Ackerman & Pritzl, 2011). However, despite the successful proof of concept demonstrated by forensic assessments, there are a number of significant barriers preventing the expansion of its use for the purpose of validity testing in child

assessments. Most notably, the MMPI-2 includes 567 items that can reportedly take multiple hours to complete (Hathaway et al., 1989). In addition to the significant burden on time, such an extensive inquiry into parental psychological characteristics would likely be viewed by many as highly unnecessary for assessing their child's symptoms and may be construed as invasive or offensive to other parents. Furthermore, as the ultimate goal is to increase the overall rate for validity measures in child assessments to levels that more closely reflect those of adult assessments, such a comprehensive and lengthy instrument may make generalization difficult across settings.

One alternative that represents a more appealing compromise would be to use a stand-alone measure of response bias, thereby eliminating the need for large item pools. One such measure that has been reliably used to identify socially desirable responders across a variety of clinical and research areas is the *Marlowe-Crowne Social Desirability Scale* (MCSDS; Crowne & Marlowe, 1960). The MCSDS is a 33-item self-report measure that utilizes a *True/False* response format where the respondent is presented with a series of low-frequency or improbable statements (Crowne & Marlowe, 1960). The measure does not include any reference to psychopathology and has exhibited high internal reliability (a = .88) and concurrent validity correlations using the MMPI-2 scales (Loo & Thorpe, 2000). A meta-analysis was conducted reviewing the use of the MCSDS across a variety of health-related settings (Van de Mortel, 2008). Overall, results revealed that adults were more likely to produce socially desirable responses when reporting on areas that might relate to their competence or to potentially stigmatizing topics (e.g., incarceration, psychiatric symptoms and diagnoses, substance abuse, levels of diet and exercise) and less likely to respond in a socially desirable manner when reporting on individually held beliefs or attitudes (Van de Mortel, 2008).

While the MCSDS potentially represents a brief and reliable alternative to the MMPI-2 in detecting defensive responders, implementation of the measure in child assessment practices may also face challenges with generalization. Specifically, the high face-validity of items within the scale may deter some parents from completing the measure as it does not directly relate to the purpose for which they sought services, the assessment of their child's psychological symptoms. In addition, while the time requirement for completion of the MCSDS is significantly less than that needed for the MMPI-2, the addition of any measure, however brief, that does not directly contribute to understanding the child's symptoms and behaviors may fail to generalize to clinical settings with significant resource restrictions for completing diagnostic assessments. Successful early acceptance and generalization of validity testing in child assessments requires researchers to overcome these challenges.

One possibility exists wherein clinicians and researchers would utilize an existing measure that both includes a way of detecting desirable responses and exhibits utility for parents during child assessments. The *Parenting Stress Index* and an updated abbreviated version, the *Parenting Stress Index/Short Form* (PSI; PSI/SF; Abidin, 1990, 1995), encapsulate both of these criteria and were specifically designed to serve as rapid screening measures for identifying stress within the parent-child relationship. The defensive responding items are embedded within a broader subscale measuring parental reports of distress in their role as a caregiver. The previously described challenges to acceptability of other validity instruments would be unlikely to apply to the use of the PSI/SF as items can be viewed by parents as directly applicable to the child assessment process. Furthermore, the MMPI-2 and MCSDS assess an individual's reports about themselves, while the PSI/SF more specifically assesses the individual's responses regarding their role as a parent and the stressors associated with that position (Abidin, 1995).

Therefore, research examining parental defensive responding may extrapolate beyond the constraints imposed by measures that solely evaluate individual characteristics whereby they do not provide the needed pattern of desirable responding regarding others required in child assessment practice.

In sum, there is inherent difficulty in measuring constructs that represent some quantity of an overt or covert characteristic during psychological assessments. This problem is exacerbated when conducting assessments with children, who traditionally were thought to be unreliable reporters. Despite significant research demonstrating that children may serve as valid and reliable sources of their own feelings and behaviors, typical practice includes gathering information from multiple sources that have produced robust low rates of agreement across informants. Prior research investigating the sources for this low concordance between raters have been far reaching with numerous studies examining specific parental characteristics (e.g., psychopathology, stress, socioeconomic status), child factors (e.g., age, gender, language development), types of assessment, reasons for referral, and severity of symptoms. Yet, despite the preponderance of research conducted in this area over the last few decades, scant attention has been given to the fundamental concept of deception within child assessments. In fact, the recent meta-analysis conducted by De Los Reyes and colleagues examined 341 correspondence studies for child assessments spanning the last 25 years of research where they identified an overwhelming number of specific predictors and pattern variations across informants (2015). However, their findings and discussion regarding the role of impression management was restricted to one paragraph cautioning clinicians and researchers about the potential minimizing influence of social anxiety on adolescent self-reports (De Los Reyes et al., 2015). There was no mention regarding the possibility that parental impression management even occurs, let alone

that it may account for differences in informant discrepancies. In fact, the authors suggested that "adult reports reliably and validly reflect patients' concerns to a greater extent than patient self-reports" (De Los Reyes et al., 2015). Furthermore, the studies previously cited within the current literature review that provided evidence for parental impression management within high stakes child forensic assessments were systematically excluded from the review by De Los Reyes and colleagues and no rationale was provided for their absence. The exclusion of these studies likely reflects a broader system of beliefs amongst child psychologists who have continuously failed to recognize and integrate validity practices with children for a variety of reasons (e.g., belief that deceptive children are easily detected, children lack incentive to deceive, children lack understanding of societal stigma). However, these views seem astonishing in light of the overwhelming adult assessment literature on detecting and correcting for invalid responses. Also, these ideas ignore the basic fact that adult informants, parents, currently operate within the assessment process. Furthermore, as typical adult assessments currently integrate some form of validity testing as standard practice, were these parents, as adults, to engage in their own psychological assessment, they would be given these measures. Why then would these same individuals not need to undergo similar requirements when asked to provide information about another person (i.e., their child), when validity testing is considered best practice in self-reports for adults? Regardless of the beliefs surrounding deception in children, the need for validity testing with adults has been well-established and as child assessment practices continue to heavily rely on parents to serve as informants, systematically excluding validity testing within this population ignores the decades of research espousing its merit.

Purpose

Validity tests are generally considered common practice in adult assessments in order to determine the veracity of individual responses. In contrast, attention to the importance of validity testing in the child assessment literature has been largely neglected despite psychologists' well-documented inability to detect invalid data without secondary measures. High rates of informant discrepancies have remained constant; however, research has failed to identify a clear pattern for understanding the factors responsible for informant discrepancies in child assessments. Whereas consequences for producing false positive data may include lowered self-esteem and stigma, an arguably greater concern is the potential for false negative data, as individuals who are truly in need of psychological services may be denied resources or treatment. Furthermore, the inconsistent standards for inclusion in research studies may lead to methodologically different conclusions across sites. Therefore, the current study was designed to apply the already well-established adult practice of using symptom validity measures to determine the contribution of respondent veracity in the child assessment process. Despite the well-documented importance of including validity testing with adult self-report measures, this study represents a novel application of these practices as applied to parents as adult-raters of their child's symptoms. By using an embedded measure of parental defensive responding, the aim of the study is to determine if higher rates of parental defensive responding will predict greater discrepancies between parent-child informant agreement across reports of the child's anxious, depressive, inattentive, and aggressive behaviors. By targeting an adult population frequently relied on to serve as informants and for whom validity testing has been well-established, the aim is for this and future studies to provide a foundation for the practical use and implementation of symptom validity measures with parents during child psychological assessments.

Hypotheses

Hypothesis 1: Low-to-moderate agreement between informants regarding child symptoms have been consistently demonstrated with rates ranging from $r = .25$ to $.28$ across hundreds of examined studies (De Los Reyes et al., 2015). Therefore, similar low-to-moderate rates of parent-child agreement are hypothesized by utilizing a parallel parent and child-report measure. In addition, differences between parent-child reports will be assessed across four child symptom constructs for which children are commonly referred for outpatient assessment and treatment (i.e., anxiety, depression, inattention, and aggression). These differences are hypothesized to represent higher parental endorsement of child symptoms than child self-endorsement across constructs.

Hypothesis 2: Discrepancies across informant agreement are hypothesized to be moderated by parental defensive responding on an embedded validity measure. Specifically, higher parental defensive responding is hypothesized to be associated with higher levels of discrepancy between parent-child informants, as high parental defensive responding will be associated with lower rates of parental symptom endorsement. Therefore, higher discrepancies will result from parents' under-reporting their child's symptoms (i.e., reporting lower overall symptoms compared to their children). Conversely, lower patterns of parental defensive responding will be associated with higher parent-child agreement across symptom constructs, as parents who respond more accurately should produce responses that are congruent with their child's self-reported symptoms.

These results are expected across each assessed symptom domain (i.e., anxiety, depression, inattention, and aggression) as previous studies assessing predictors of informant disagreement have not determined a consistent pattern for describing informant discrepancies.

While the research in this area remains scarce, this pattern of responses is theoretical and may result for a variety of reasons (e.g., self-referral due to the child's recognition of service need rather than parental referral, parental recognition that a problem exists but difficulty endorsing impairment for fear of stigma, parental fear that reporting symptoms in their child will reflect negatively on them and their abilities as a parent).

Hypothesis 3: As the proposed measure for detecting defensive response patterns is embedded within an existing measure of parental stress where lower item endorsement is considered significant, statistical analyses for assessing the pattern of discrepancy are hypothesized to produce a curvilinear relationship across all assessed symptom domains (i.e., anxiety, depression, inattention, and aggression). Specifically, as parental defensive responding is hypothesized to correlate with lower levels of parent-endorsed child symptoms, a curvilinear pattern of lower agreement at higher levels of parental defensive responding and higher agreement at lower levels of parental defensive responding is hypothesized. Laird and LaFleur (2016) first proposed using polynomial regression techniques for parent-child informant discrepancies as it allows agreement to serve as the dependent variable and interaction terms can be used to measure levels of predictor correlations. Therefore, by evaluating the linear, quadratic, and interaction coefficient terms between parallel parent-child symptom domains (i.e., anxiety, depression, inattention, and aggression) a hierarchical polynomial regression is hypothesized to serve as a reliable method for measuring patterns of informant discrepancies across child symptom constructs.

Method

Participants

This study made use of an existing and ongoing sample of families seeking psychoeducational assessments and/or treatment at the Psychological Services Center (PSC) on Louisiana State University's main campus. All participants were either self-referred in response to clinic publicity or referred by school personnel and other mental health or community professionals. Participants were drawn from the database that currently contains more than 400 individuals, and included only if they had completed at least 85% of each of the study measures.

The final sample of participants included a total of 220 parent-child dyads. The children were 42.7% female and ranged in age from 11-17 years ($M = 12.52$; $SD = 1.43$). This restriction in age range is a deviation from that found within the overall database and is a product of the selection of the specific measure utilized for assessing child-reported symptoms. Participants self-identified as White/Caucasian (86.4%; $n=190$), Black/African American (8.2%; $n=18$), Asian (2.7%; $n=6$), and Hispanic (1.4%; $n=3$). Three participants (1.4%, $n=3$) either chose not to identify or the information was missing from the database. No participant was excluded on the basis of gender, race, ethnicity, or sexual orientation. Missing values were replaced for two participants on one measure using within-subject means. This is a conservative replacement procedure as within-subject mean replacement may be utilized effectively for up to 20% data loss (Downey & King, 1998).

Measures

In order to assess a defensive response pattern for this study, the *Parenting Stress Index/Short Form* (PSI/SF; Abidin, 1990, 1995) was completed by parents. The PSI/SF is a 36-item parent-report measure designed to target families in need of assistance due to stress within

the parent-child relationship. The measure takes approximately ten minutes to administer and was included in the standard psychoeducational battery given to all individuals seeking services at the PSC. The embedded measure of defensive responding included in the measure was modeled after the *Marlowe-Crowne Social Desirability Scale* (MCSDS; Crowne & Marlowe, 1960), which is generally considered to be the most accurate stand-alone measure for detecting social desirability in respondents (Lambert, Arbuckle, & Holden, 2016). The Defensive Responding Index on the PSI/SF was designed to assess parental attempts to present themselves in a favorable manner to the clinician (Castaldi, 1988; Lafiosca & Loyd, 1986). The Defensive Responding Index comprises seven items embedded within the Parental Distress subscale. Scale items utilize the same techniques that have been proven most successful at detecting non-credible responses in adult self-report measurements. Specifically, the PSI/SF was developed using ideal items and inverted responses. The Defensive Responding Index score is calculated by summing the raw scores across the seven scale items so that lower scores are suggestive of clinical significance while higher scores represent accurate parental responding. For the current study, internal consistency across participants was calculated as good for the Defensive Responding Index ($\alpha = .93$).

The *Child Behavior Checklist for Ages 6 to 18* (CBCL/6-18; Achenbach & Rescorla, 2001) is a self-report measure completed by parents. It was designed to assess a broad range of children's competencies, adaptive functioning, and problems. The measure produces three broadband indices (Internalizing, Externalizing, Total Problems), eight syndrome subscales (Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, Aggressive Behavior), and three adaptive competency scales (Activities, Social, School). Caregivers are asked to rate each item

on a three-point Likert scale (0 = *not true*, 1 = *somewhat* or *sometimes true*, 2 = *very true* or *often true*). Syndrome subscales are computed by totaling the raw scores for each associated item and higher scores indicate clinical difficulties. Internal consistency for the three broadband indices was calculated as >.90 and mean alphas for the subscales =.84 (Achenbach & Rescorla, 2001). For the current study, four syndrome subscales were used (i.e., Anxious/Depressed, Withdrawn/Depressed, Attention Problems, and Aggressive Behavior). Research has shown significant associations between subscale scores on the CBCL and specific diagnostic categories (Achenbach & Rescorla, 2001; Kasius, Ferdinand, van den Berg, & Verhulst, 1997). Correlations have been demonstrated between the Anxious/Depressed subscale and anxiety disorder diagnoses (r = .51), the Withdrawn/Depressed subscale and depressive diagnoses (r = .49), the Attention Problems subscale and ADHD (r = .80), and the Aggressive Behavior subscale and oppositional defiant disorder (r = .64) using a clinical interview evaluation format (Achenbach & Rescorla, 2001). In addition, the Anxious/Depressed subscale has not been found to significantly correlate with depressive diagnoses. Therefore, for ease of understanding across symptom constructs, subsequent references to the syndrome subscales used throughout this study will be simplified to reflect their underlying constructs (i.e., Anxious/Depressed= anxiety, Withdrawn/Depressed= depression, Attention Problems= inattention, Aggressive Behavior= aggression). For the current study, internal consistency across participants was calculated as good (α = .91 for Anxious/Depressed, α = .89 for Withdrawn/Depressed, α = .86 for Attention Problems, and α = .83 for Aggressive Behavior).

The *Youth Self-Report* (YSR/11-18; Achenbach & Rescorla, 2001) is a 112-item self-report questionnaire completed by the youth. It was designed to parallel the indices present on the parent-report version of the CBCL. Specifically, the YSR assesses the same broad range of

youth difficulties, adaptive functioning, and competencies and produces dimensional scores across the same eight syndrome subscales (Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, Aggressive Behavior). In the current study, the same four syndrome subscales were used to determine informant discrepancies between parent and child ratings (i.e., Anxious/Depressed, Withdrawn/Depressed, Attention Problems, and Aggressive Behavior). For the current study, internal consistency was determined as good ($\alpha = .93$ for Anxious/Depressed, $\alpha = .92$ for Withdrawn/Depressed, $\alpha = .84$ for Attention Problems, and $\alpha = .86$ for Aggressive Behavior).

Procedure

Participants were seen through the PSC at Louisiana State University. Primary caregivers and children served as informants for the study. Parental informed consent and child assent were obtained at the time of the assessment. Each parent-child dyad was administered a standardized battery of psychoeducational and research measures. IRB approval for the database was recently renewed until April 2018 and has been ongoing since 2006. All measures were administered by graduate students in clinical psychology under the direction of a licensed supervisor. All students were trained to administer the battery of measures according to standardized protocols provided within each test's respective technical manual. Supervision was provided through ongoing practicum to ensure maintenance of consistency and fidelity of test administration.

Results

Power

In order to determine the necessary sample size for adequate power, an a priori G*Power version 3.1 analysis was utilized (Faul, Erdfelder, Buchner, & Lang, 2009). Power analysis was conducted for the polynomial regression analyses as they were considered the largest power burden for the study. In order to determine the minimum number of participants needed for the study, alpha was set to .05 and power was set to .80 (Cohen, 1988). A relatively small effect size of $f^2 = 0.08$ was selected, as the power to detect small to medium effects within polynomial regression equations has been previously demonstrated (Laird & De Los Reyes, 2013; Sibley, Campez, & Raiker, 2017). The power analysis recommended a total sample size of $n = 124$. As the number of participants included for analyses in the current study is $n = 220$, this indicates that the analyses were sufficiently powered to detect all statistics of interest. All other statistical analyses were conducted using SPSS version 23.

Descriptive Statistics

Parent and child raw scores from the four symptom domains across parallel measures (i.e., CBCL and YSR; Anxious/Depressed, Withdrawn/Depressed, Attention Problems, Aggressive Problems) were converted to t-scores with a mean of 50 and a standard deviation of 10. Paired sample t-tests were calculated between parent and child reports across the four parallel symptom domains to determine whether the difference between reports was statistically significant. Table 3 provides the results of these analyses.

Table 3.

Congruence between Parent and Child Reports across Symptom Domains

	Parent Report	Child Report	Difference of Means	
	M (*SD*)	*M* (*SD*)	T	*p*
Anxiety	58.06 (8.72)	55.77 (7.34)	3.53	.001
Depression	57.09 (8.48)	55.20 (6.71)	3.28	.001
Inattention	64.13 (10.12)	60.55 (9.42)	4.66	<.001
Aggression	56.80 (7.85)	55.10 (7.14)	3.13	.002

Overall, parents reported higher rates of anxiety, depression, inattention, and aggressive symptoms than did their paired child. Parent's reported a statistically significant increase in their child's anxiety symptoms compared to child self-report ratings of their anxiety symptoms ($M =$ 2.286 points, $SE = .648$, $t(219) = 3.531$, $p = .001$, $d = 0.24$). Parent's reported a statistically significant increase in their child's depression symptoms compared to child self-report ratings of their depression symptoms ($M = 1.891$ points, $SE = .577$, $t(219) = 3.275$, $p = .001$, $d = 0.22$). Parent's reported a statistically significant increase in their child's inattention symptoms compared to child self-report ratings of their inattention symptoms ($M = 3.582$ points, $SE = .769$, $t(219) = 4.657$, $p < .001$, $d = 0.31$). Parent's reported a statistically significant increase in their child's aggression symptoms compared to child self-report ratings of their aggression symptoms ($M = 1.695$ points, $SE = .542$, $t(219) = 3.128$, $p = .002$, $d = 0.21$). Correlations between study variables along with means and standard deviations were computed and are presented in Table 4.

Table 4.
Correlations, Means, and Standard Deviations for Moderators and Key Variables

	P ANX	C ANX	P DEP	C DEP	P INT	C INT	P AGG	C AGG	P DR	Mean (SD)
P ANX	--									58.06 (8.72)
C ANX	.294**	--								55.77 (7.34)
P DEP	.502**	.147*	--							57.09 (8.48)
C DEP	.171*	.584**	.384**	--						55.20 (6.71)
P INT	.240**	.001	.374**	.039	--					64.13 (10.12)
C INT	.101	.359**	.090	.278**	.320**	--				60.55 (9.42)
P AGG	.449**	.129	.406**	.203**	.467**	.227**	--			56.80 (7.85)
C AGG	.116	.472**	.132	.504**	.109	.510**	.427**	--		55.10 (7.14)
P DR	.238**	-.022	.183**	.067	.178**	.013	.286**	.086	--	15.76 (6.09)
C Age	-.072	-.027	.036	-.011	-.052	-.019	-.094	-.116	-.008	12.42 (1.43)

Note. ANX= Anxiety; DEP= Depression; INT= Inattention; AGG= Aggression; DR=Defensive Responding
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

No significant correlations were identified with child age on any of the parent or child report scales (all $p > .05$). There was a positive correlation between parent-child reports across parallel symptom scores for all assessed domains (Anxiety $r = .294$, $p < .001$; Depression $r = .384$, $p < .001$; Inattention $r = .320$, $p < .001$; Aggression $r = .427$, $p < .001$). Furthermore, there was a small positive correlation between parental defensive responding and each parent reported symptom variable (Parent Anxiety Rating $r = .238$, $p < .001$; Parent Depression Rating $r = .183$, $p < .001$; Parent Inattention Rating $r = .178$, $p < .001$; Parent Aggression Rating $r = .286$, $p < .001$).

Assumptions

There are six assumptions that must be tested prior to conducting a multiple linear regression analysis and polynomial regression requires five of the same assumptions with the exception of requiring that a linear relationship be demonstrated (Berry, 1993). The first assumption requires a continuous dependent variable and two or more independent variables. This assumption was not violated as all variables were measured at the continuous level. Second, the sample was analyzed for significant outliers. In order to test for significant outliers, standardized residuals were calculated for each study variable and a criteria value of ±3 standard deviations was used as representative criteria (Cook & Weisberg, 1982). This procedure did not identify any significant outliers and no influential points were identified using a criteria of Cook's Distance values above 1.0 (Cook & Weisberg, 1982). Third, the assumption of multicollinearity was tested by inspecting the correlation coefficients for all items. As tolerance values were greater than 0.1, there was no evidence of multicollinearity. Fourth, the Durbin-Watson statistic was used to test the assumption of independence of observations. All study variables showed independence of residuals using the Durbin-Watson statistic with values ranging from 1.74 to 2.05. The fifth assumption of *homoscedasticity*, variances along the line of best fit remain constant across the line, was assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. As variances remained relatively constant across the line of best fit, this assumption was not violated. Finally the sixth assumption requires that all residuals (errors) are approximately normally distributed. The assumption of normality was met by assessing a produced Q-Q plot for each regression model.

Polynomial Regression Analyses

The polynomial regression model in this study utilizes the equation technique developed by Laird and LaFleur for assessing informant discrepancies as the dependent variable (2016). While polynomial regression allows for any number of increasingly higher order terms to be examined (i.e., X^2, X^3…X^{14}, etc.) there is considerable evidence suggesting that the types of constructs and measurement instruments used in psychology rarely exceed the quadratic level (Cohen & Cohen, 1983; Cortina, 1993; Dawson, 2014). Specifically, raising the power of a predictor variable beyond the second-order quadratic term would represent the effects of that variable at increasingly more extreme levels that are not interpretable or meaningful for the majority of psychological constructs. Therefore, the variable terms used in the current study do not exceed the quadratic level. The model for testing parental defensive responding (D) as a predictor of child (C) and parent (P) informant discrepancy is as follows:

$$C = b_0 + b_1P + b_2P^2 + b_3D + b_4D^2 + b_5PD + b_6P^2D + b_7PD^2 + e$$

The predictor of parent-child informant discrepancy (i.e., parent rated defensive responding) was examined across four separate hierarchical polynomial regression analyses with child-reported symptom severity (i.e., anxiety, depression, inattention, aggression) representing the dependent variable and parent-reported symptom severity representing the independent variable paired with the same child symptom construct. Therefore, the linear coefficient (b_1) reflects the effects of parent-reported symptom severity as a predictor of child-reported symptom severity and the linear coefficient (b_3) reflects the effects of parent-reported defensive responding as a predictor of child-reported symptom severity.

The coefficient (b_5) represents the linear interaction effect between parent-reported symptom severity and parent-reported defensive responding as a predictor of child-reported symptom severity. Therefore, the interaction term for the current model examines the extent to which parent-child congruence in their symptom reports is moderated by parental defensive responding.

The quadratic terms in the model represent nonlinear relationships. The coefficient (b_2) reflects the quadratic effect of parent-reported symptoms as a predictor of child self-reported symptoms. Specifically, a significant quadratic effect would suggest that prediction of child self-reported symptoms varies as a function of the level of parent reported symptom severity. The coefficient (b_4) reflects the quadratic effect of parental defensive responding on child self-reported symptoms where prediction of child self-reported symptoms varies as a function of the level of parental defensive responding. While this term seems unlikely to predict the outcome variable based on a theoretical understanding of the relationship between the constructs, inclusion of both quadratic terms in the model is required when also examining quadratic interaction effects (Ganzach, 1997; LeBreton, Tonidandel, & Krasikova, 2013). Exclusion of a quadratic term when also assessing the quadratic interaction coefficients has been shown to confound interpretation of the variance explained by the quadratic effects by inflating the significance of the regression weights due to the unexamined component of the interaction (Cortina, 1993).

The coefficients (b_6) and (b_7) represent the quadratic interaction effects between parent reported child symptoms and parental defensive responding. Specifically, (b_6) reflects the quadratic interaction effect where prediction of child self-reported symptoms varies as a function of the level of parent reported symptom severity that is moderated by parental defensive

54

responding. In contrast, (b_7) reflects the quadratic interaction effect whereby prediction of child self-reported symptoms varies as a function of the level of parental defensive responding that is moderated by parent reported symptom severity.

Finally, the coefficient (b_0) represents the intercept for the polynomial regression equation and the term (e) denotes the residual variance that is unexplained by the polynomial regression equation. Therefore, each symptom regression model initially included seven terms: parent reported symptom, parent reported symptom squared, parent reported defensive responding, parent reported defensive responding squared, a linear interaction term of parent reported symptom multiplied by parent reported defensive responding, and two quadratic interaction terms of parent reported symptom squared multiplied by parent reported defensive responding and parent reported symptom multiplied by parent reported defensive responding squared.

Prior to conducting data analyses to test the equation model above, parent reported defensive responding and parent reported child symptoms across all four constructs (i.e., anxiety, depression, inattention, aggression) were mean-centered according to established standards (Dawson, 2014; Kleinbaum et al., 2013). These mean-centered values were then used to compute new variables to represent those found in the equation model. Specifically, the mean-centered defensive responding variable was multiplied by itself to create the quadratic term (D^2) and the mean-centered variable for each parent reported child symptom construct was separately used to create the quadratic terms (P^2) for their respective models. A linear interaction term for each model (PD) was computed by multiplying the mean-centered defensive responding variable by the mean-centered parent reported child symptom associated with each respective symptom construct model. The quadratic interaction term (P^2D) was computed by multiplying the mean-

centered parent reported child symptom variable by itself and by the mean-centered defensive responding variable. Finally, the quadratic interaction term (PD^2) was computed by multiplying the mean-centered parent reported child symptom variable by the mean-centered defensive responding variable and by itself.

Following the computation of these additional variables, four separate polynomial regression analyses were conducted, one for each child symptom construct. The polynomial regression analyses were conducted hierarchically with two separate steps. The initial step contained the first 5 independent variables from the equation (i.e., parent reported symptom, parent reported symptom squared, parent reported defensive responding, parent reported defensive responding squared, parent reported symptom multiplied by parent reported defensive responding). The second step included the final two independent variables from the equation (i.e., parent reported symptom squared multiplied by parent reported defensive responding and parent reported symptom multiplied by parent reported defensive responding squared). The second step was retained only if there was significant improvement in the fit of the regression model, which was assessed by evaluating the change in R^2.

Congruence and plotting predicted values. Across the four polynomial regression models (i.e., anxiety, depression, inattention, aggression), one quadratic interaction term was significant. The quadratic interaction term showed that parent reported child aggression symptoms interacted with parental defensive responding and parental defensive responding to predict child self-reported aggression symptoms. All significant terms were plotted by calculating 1 *SD* above and below the mean levels of the moderator (Aiken & West, 1991; Cohen & Cohen, 1983; Cohen, Cohen, West, & Aiken, 2003). Specifically, plotted predicted values were determined by fitting the polynomial regression equation with parental defensive

responding at +1 *SD* above the mean (i.e., lower levels of parental defensive responding) and -1

SD below the mean (i.e., higher levels of parental defensive responding) to calculate the

predicted range of child reported symptoms across parent reported symptoms.

The simple slopes were calculated and plotted using the PROCESS add-on version 2.16

for SPSS and are shown in Table 5 (Hayes, 2013). Notably, this technique of plotting points

above and below mean levels does not create a dichotomous term; rather, it produces high and

low scores along the continuous slope so that prediction values may be calculated across the full

range of the equation. Perfect parent-child agreement for all symptom models is visually

represented to aid in interpretability across figures as a light gray solid diagonal line representing

where Y = X (i.e. from the origin in the bottom-left to the top-right corner).

Table 5.
Means for Slopes at Each ± 1 SD Level of Parent Reported Defensive Responding

	1 SD Below Mean DR		1 SD Above Mean DR	
	Child Report (-1 SD Parent)	Child Report (+1 SD Parent)	Child Report (-1 SD Parent)	Child Report (+1 SD Parent)
Anxiety	50.44 (50.00)	60.05 (66.78)	50.46 (50.00)	57.82 (66.78)
Depression	50.91 (50.00)	61.18 (65.57)	50.90 (50.00)	57.64 (65.57)
Inattention	58.19 (54.01)	62.01 (74.25)	55.68 (54.01)	66.35 (74.25)
Aggression	54.42 (50.00)	59.39 (64.65)	51.99 (50.00)	56.61 (64.65)

Note. DR = Defensive Responding; Values of parental defensive responding at -1 SD below the
mean (9.67 raw score) reflect higher levels of the construct and values at +1 SD above the mean
(21.85 raw score) reflect lower levels of parent reported defensive responding.

The area of the graph above this diagonal represents higher child self-reported symptoms

compared to their parent's reports. Conversely, the area of the graph below the divider represents

higher parent reported child symptoms compared to child self-reports. By examining the plotted

predicted values in relation to this dividing line, informant agreement can be interpreted at

varying levels of congruence based on the distance of the plotted value from the line. Notably,

the inclusion of quadratic effects in the regression model has been shown to produce curved lines when plotting the predicted values even when only linear interactions are significant (Cohen & Cohen, 1983).

Anxiety Model. The overall polynomial regression model (Table 6) for predicting child self-reported anxiety symptoms was statistically significant, $R^2 = .125$, $F(7, 212) = 4.341$, $p < .001$; adjusted $R^2 = .096$. The addition of higher order quadratic interaction terms in the prediction (Step 2) was not statistically significant, R^2 increase of .007, $F(2, 212) = 0.848$, $p > .05$. Therefore, based on prior recommendations, the two higher order quadratic interaction terms (b_6P^2D) and (b_7PD^2) were removed from the prediction equation (Edwards, 1994; Laird & De Los Reyes, 2013; Laird & LaFleur, 2016).

The simple linear effect of parent reported child anxiety was statistically significant ($p < .001$) and the simple linear effect of parent reported defensive responding was statistically significant ($p < .05$). In other words, parent-reported child anxiety and parent-reported defensive responding both, separately, significantly predicted child self-reported anxiety symptoms: more parent-reported anxiety symptoms were predictive of more child self-reported anxiety symptoms, and more parent-reported defensive responding predicted less child self-reported anxiety. The final model tested for the anxiety symptom domain was reduced to the following equation:

$$C = b_0 + b_1P + b_2P^2 + b_3D + b_4D^2 + b_5PD + e$$

As no significant interaction or quadratic effects were found for this model, the graph (Figure 1) of predicted values of child reported self-anxiety symptoms represents the simple linear effect of parent reported child anxiety symptoms at 1 *SD* above and below the mean for parental defensive responding. The predicted values were calculated by substituting the

regression coefficients produced from the data into the final polynomial equation model.

Notably, the inclusion of quadratic effect terms in the equation model resulted in curved lines for

the plotted values.

Table 6.
Polynomial Regression Analyses Predicting Child Reports of Anxiety Symptoms from Parents
Reports of Anxiety Symptoms and Parents Defensive Responding

Parameter	Anxiety		
	B	SE	p
Step 1			
Parent report	3.318	.670	<.001
Parent report squared	-.779	.431	.072
Parent DR	-1.128	.560	.045
Parent DR squared	.526	.358	.143
Parent report x Parent DR	-.015	.553	.978
Step 2			
Parent report squared x Parent DR*			
Parent report x Parent DR squared*			
Step 1 Model R^2	.118	6.976	<.001
Step 2 Model R^2	.125	6.981	.430
R^2 Change	.007		

Note. DR= Defensive Responding; B= unstandardized regression coefficient; SE= standard error
of the coefficient;* higher-order quadratic interaction terms were removed from final model
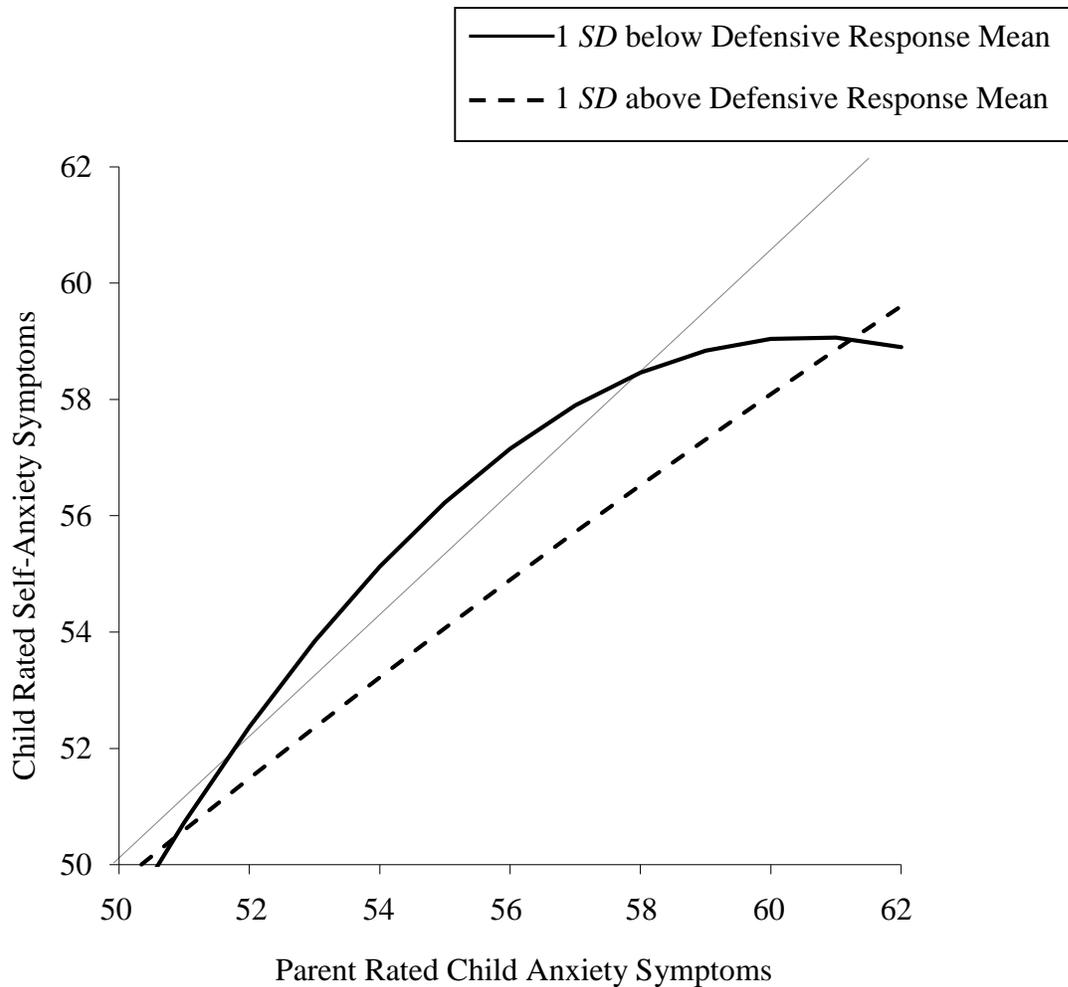
Figure 1. Predicted Child Reported Anxiety Symptoms as a Function of Parent Reported Child Anxiety Symptoms and Parent Reported Defensive Responding

As shown in Figure 1, the linear effect of parent reported child anxiety symptoms is relatively constant at higher levels of parental defensive responding (-1 *SD* below Mean [reflecting higher defensive responding] $b = 2.8$, $SE = 1.25$, p = .025) and lower levels of parental defensive responding (+1 *SD* above Mean [reflecting lower defensive responding] $b = 2.27$, $SE = 0.66$, $p = <.001$). The lighter diagonal line across the figure was provided to illustrate perfect parent-child symptom congruence. Visual inspection of the graph shows that the effect of parent reported anxiety symptoms on child self-reported anxiety symptoms was negative and concave. A concave effect represents a relationship where parent reported child anxiety

symptoms increase as child reported symptoms increase but only to a certain level of symptom endorsement. After reaching the peak of the curve, any further increases in parent reported anxiety symptoms will yield lower levels of child self-reported symptoms.

Depression Model. The overall polynomial regression model (Table 7) for predicting child self-reported depression symptom was statistically significant, $R^2 = .174$, $F(7, 212) = 6.362$, $p < .001$; adjusted $R^2 = .146$. The addition of higher order quadratic interaction terms in the prediction (Model 2) was not statistically significant, R^2 increase of .017, $F(2, 212) = 2.121$, $p > .05$.

Table 7.
Polynomial Regression Analyses Predicting Child Reports of Depression Symptoms from Parents Reports of Depression Symptoms and Parents Defensive Responding

Parameter	Depression		
	B	*SE*	*p*
Step 1			
Parent report	3.284	.653	<.001
Parent report squared	-.346	.261	.185
Parent DR	-.314	.489	.521
Parent DR squared	.308	.318	.334
Parent report x Parent DR	-.136	.430	.752
Step 2			
Parent report squared x Parent DR*			
Parent report x Parent DR squared*			
Step 1 Model R^2	.157	6.234	<.001
Step 2 Model R^2	.174	6.202	.122
R^2 Change	.017		

Note. DR= Defensive Responding; B= unstandardized regression coefficient; SE= standard error of the coefficient; * higher-order quadratic interaction terms were removed from the final model

Therefore, based on prior recommendations, the two higher order quadratic interaction terms (b_6P^2D) and (b_7PD^2) were removed from the prediction equation (Edwards, 1994; Laird & De Los Reyes, 2013; Laird & LaFleur, 2016). The final model tested for the depression symptom domain was reduced to the following equation: $C = b_0 + b_1P + b_2P^2 + b_3D + b_4D^2 + b_5PD + e$.

61

The simple linear effect of parent reported child depression was statistically significant (p <.001). In other words, parent reported child depression significantly predicted child self-reported depression: more parent reported depression symptoms were predictive of more child self-reported depression symptoms. As no significant interaction or quadratic effects were found for this model, the graph (Figure 2) of predicted values of child reported self-depression symptoms represents the simple linear effect of parent reported child depression symptoms at 1 *SD* above and below the mean for parental defensive responding.
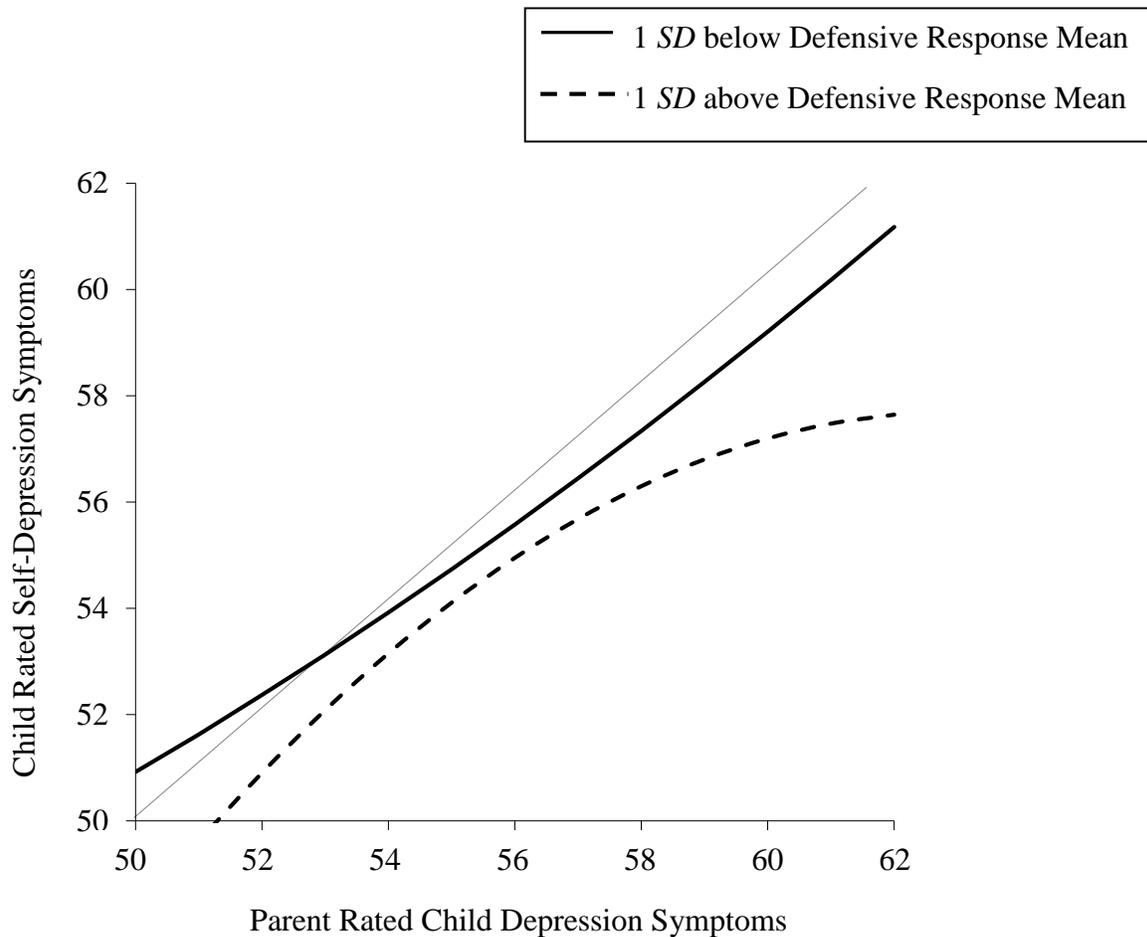

Figure 2. Predicted Child Reported Depression Symptoms as a Function of Parent Reported Child Depression Symptoms and Parent Reported Defensive Responding

The predicted values were calculated by substituting the regression coefficients produced from the data into the final polynomial equation model. Notably, the inclusion of quadratic effect terms in the equation model resulted in curved lines for the plotted values. The simple linear effect of parent reported child depression symptoms remains relatively constant across levels of parental defensive responding (+1 *SD* above Mean [reflecting lower defensive responding] $b =$ .29, *SE* = .08, *p* < .001; -1 *SD* below Mean [reflecting higher defensive responding] $b =$ .32, *SE* = .09, *p* = .001). The lighter diagonal line across the figure was provided to illustrate perfect parent-child symptom congruence.

Inattention Model. The overall polynomial regression model (Table 8) for predicting child self-reported inattention symptoms was statistically significant, $R^2 = .126$, $F(7, 212) = 4.354$, *p* < .001; adjusted $R^2 = .097$. The addition of higher order quadratic interaction terms in the prediction (Step 2) was not statistically significant, R^2 increase of .007, $F(2, 212) = 0.885$, *p* > .05. Therefore, based on prior recommendations, the two higher order quadratic interaction terms (b_6P^2D) and (b_7PD^2) were removed from the prediction equation (Edwards, 1994; Laird & De Los Reyes, 2013; Laird & LaFleur, 2016). The final model tested for the inattention symptom domain was reduced to the following equation:

$$C = b_0 + b_1P + b_2P^2 + b_3D + b_4D^2 + b_5PD + e$$

The simple linear effect of parent reported child inattention was statistically significant (p <.001). In other words, parent reported child inattention significantly predicted child self-reported inattention: more parent reported inattention symptoms were predictive of more child self-reported inattention symptoms.

Table 8.

Polynomial Regression Analyses Predicting Child Reports of Inattention Symptoms from Parents Reports of Inattention Symptoms and Parents Defensive Responding

	Inattention		
Parameter	B	SE	p
Step 1			
Parent report	3.027	.699	<.001
Parent report squared	-.119	.532	.823
Parent DR	-.168	.705	.812
Parent DR squared	-.055	.460	.904
Parent report x Parent DR	1.285	.703	.069
Step 2			
Parent report squared x Parent DR*			
Parent report x Parent DR squared*			
Step 1 Model R^2	.118	8.949	<.001
Step 2 Model R^2	.126	8.954	.414
R^2 Change	.007		

Note. DR= Defensive Responding; B= unstandardized regression coefficient; SE= standard error of the coefficient * higher-order quadratic interaction terms were removed from the final model

As no significant interaction or quadratic effects were found for this model, the graph (Figure 3) of predicted values of child reported self-inattention symptoms represents the simple linear effect of parent reported child inattention symptoms at 1 *SD* above and below the mean for parental defensive responding. The predicted values were calculated by substituting the regression coefficients produced from the data into the final polynomial equation model. Notably, the inclusion of quadratic effect terms in the equation model resulted in curved lines for the plotted values. As shown in Figure 3, the linear effect of parent reported child inattention symptoms is significantly stronger at lower levels of parental defensive responding (+1 *SD* above Mean [reflecting lower defensive responding] $b = 4.31$, $SE = .77$, $p < .001$) than at higher levels of parental defensive responding (-1 *SD* below Mean [reflecting higher defensive responding] $b = 1.74$, $SE = 1.03$, $p = .091$).
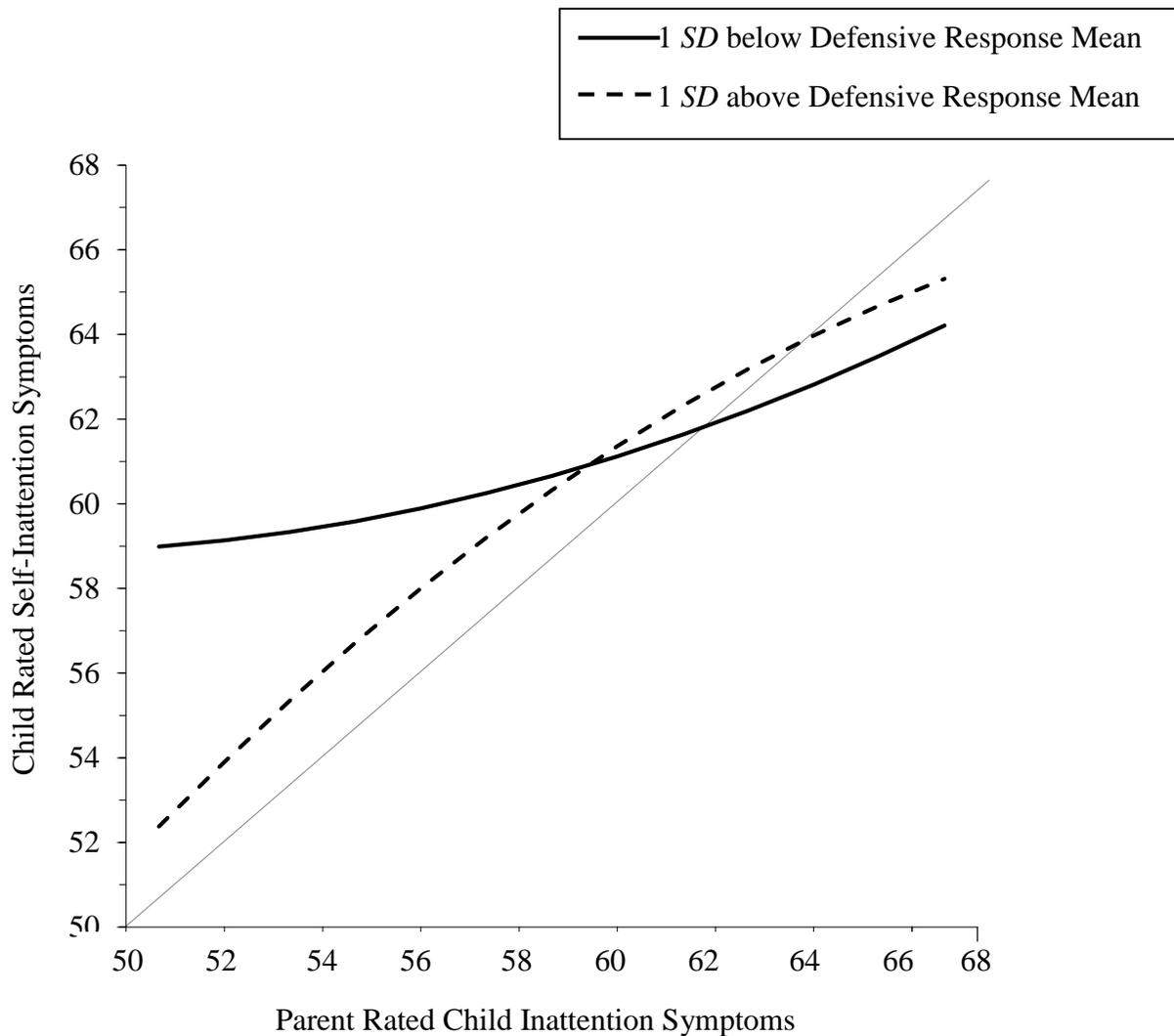
Figure 3. Predicted Child Reported Inattention Symptoms as a Function of Parent Reported Child Inattention Symptoms and Parent Reported Defensive Responding

The plotted predicted values show greater parent-child congruence for reports of inattention (near perfect agreement with the light gray line representing; Y = X) when parental defensive responding was lower (+1 *SD* above Mean; reflecting lower defensive responding). In contrast, the predicted values show when parental defensive responding was higher (-1 *SD* below Mean; reflecting higher defensive responding) informant agreement is significantly lower.

Aggression Model. The overall polynomial regression model (Table 9) for predicting

child self-reported aggression symptoms was statistically significant, $R^2 = .243$, $F(7, 212) =$

9.713, $p < .001$; adjusted $R^2 = .218$. The addition of higher order quadratic interaction terms in

the prediction (Step 2) led to a statistically significant increase in R^2 of .028, $F(2, 212) = 3.917$, p

$= .021$. Therefore, the regression equation model used for predicting child aggression symptoms

retained both steps with all originally defined coefficients:

$$C = b_0 + b_1P + b_2P^2 + b_3D + b_4D^2 + b_5PD + b_6P^2D + b_7PD^2 + e$$

Plotted predicted values were calculated by substituting the regression coefficients

produced from the data into the regression equation model. As shown in Figure 4, there is a

significant quadratic interaction effect between parental defensive responding (squared) and

parent reported aggression symptoms on predicting child self-reported aggression symptoms.

Table 9.
Polynomial Regression Analyses Predicting Child Reports of Aggression Symptoms from
Parents Reports of Aggression Symptoms and Parents Defensive Responding

Parameter	Aggression		
	B	SE	p
Step 1			
Parent report	1.597	.688	.021
Parent report squared	.732	.339	.032
Parent DR	-.251	.627	.689
Parent DR squared	.547	.331	.100
Parent report x Parent DR	-.061	.723	.932
Step 2			
Parent report squared x Parent DR	-.468	.335	.163
Parent report x Parent DR squared	.907	.353	.011
Step 1 Model R^2	.215	6.397	<.001
Step 2 Model R^2	.243	6.312	.021
R^2 Change	.028		

Note. DR= Defensive Responding; B= unstandardized regression coefficient; SE= standard error
of the coefficient

A significant quadratic interaction suggests that the linear term is moderating the steepness of the bend (i.e., the instantaneous rate of change in the curvilinearity) of the curve in predicting the outcome variable. When applied to the current model, the quadratic effect of parent reported child aggression symptoms is stronger at higher levels of parental defensive responding (-1 *SD* below Mean; [reflecting higher defensive responding] $b = .02$, $SE = .36$, $p = .03$) than at lower levels of parental defensive responding (+1 *SD* above Mean [reflecting lower defensive responding] ($b = .01$, $SE = .36$, $p = .20$).
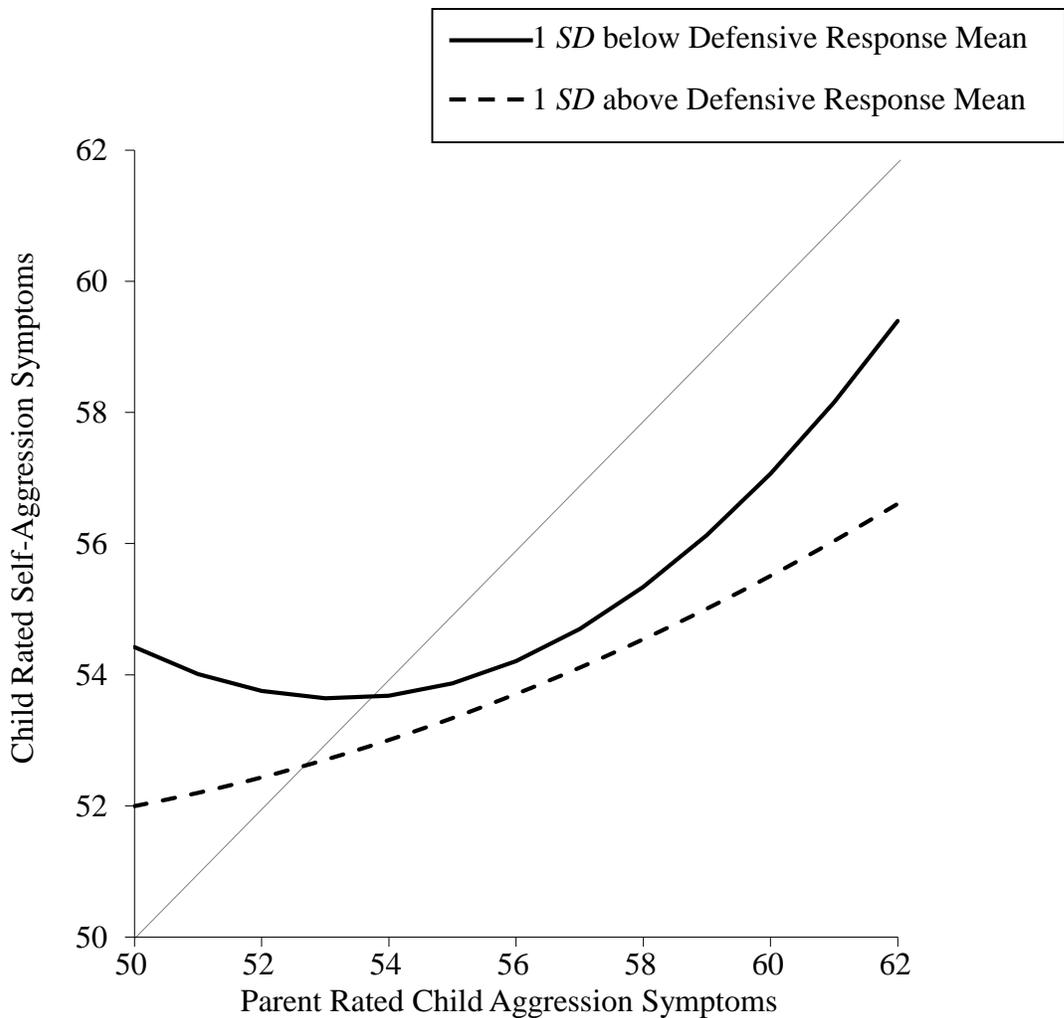

Figure 4. Predicted Child Reported Aggression Symptoms as a Function of Parent Reported Child Aggression Symptoms and Parent Reported Defensive Responding

The plotted predicted values show more congruence in parent-child reported aggression symptoms at lower levels of parental defensive responding. More specifically, the greatest levels of incongruence were associated with the significant interaction between higher levels of parental defensive responding and different values of parent reported aggression ratings. As can be seen in the graph, the steepness of the curve in parental defensive responding was moderated by the level of severity in parent reported aggression symptoms, where the highest levels of congruence between parent-child reported aggression symptoms was found around average levels of symptom severity. In addition, the interaction between higher levels of parental defensive responding and both higher and lower parent reported symptom severity was predictive of the highest levels of incongruence with child reported aggression symptoms.

Discussion

Multi-informant child assessment practices allow for additional perspectives on a child's functioning and behavior across multiple settings and contexts; however, this also leads to more complicated integration requirements when information differs across informants. Validity testing represents a well-established process for reducing error in adult assessments but has yet to be utilized in child assessment practices. In the current study, an embedded measure of defensive responding was applied to a sample of parent-child dyads to investigate its usefulness in predicting informant discrepancies. Four polynomial regression models were tested across symptom domains for which children are commonly referred for treatment.

Informant Correlations

Low to moderate rates of agreement between parent-child ratings were hypothesized across symptom domains based on previous meta-analytic studies. Current findings partially supported this hypothesis as moderate rates of agreement were found across constructs; however,

agreement was higher in the current study than anticipated. While correlations between parent-child anxiety ratings were similar to meta-analytic results (De Los Reyes et al., 2015), correlations between the other symptom constructs, particularly aggression, were much higher. There have been a significant number of studies assessing factors that may influence agreement rates; however, no one explanation exists for these findings. One possible rationale for the current sample may be lower referral for aggressive behaviors within the clinic where the assessments were conducted. As an outpatient training clinic on a university campus, the overall rate of referral for highly aggressive behavior may differ from previously conducted studies. Therefore, agreement may have been higher based solely on the lower base rate of the behavior.

The hypothesis that parents would endorse higher rates of symptom severity across constructs as compared to their children was supported overall in the current study. Parents endorsed significantly higher (all p< .005) symptom severity across the anxiety, depression, inattention, and aggression domains as compared to child reports. Theoretically, this is consistent with the idea that children rarely engage in treatment seeking behavior; rather, parents are responsible for seeking services.

Polynomial Regression Analyses

Utilizing a polynomial equation model, the current study hypothesized that varying levels of parental defensive responding would moderate the prediction of discrepancies between parent-child reports of the child's psychological symptoms. This was examined by evaluating the linear, quadratic, and interaction coefficients across four separate symptom models. Overall, results indicated that an embedded measure of parental defensive responding was able to predict some of the discrepancy between parent-child reports for child symptoms; however, results were significantly different across symptom constructs.

Symptom models. Hypotheses predicted that a significant interaction effect would produce higher rates of informant discrepancy across parent-child dyads. The results from the models of internalizing symptoms did not support these expectations, as neither model resulted in a significant interaction effect. Results from the anxiety polynomial regression model resulted in significant simple linear effects for both parent reported anxiety and parental defensive responding; however, the interaction term for parental defensive responding by parent reported anxiety symptoms was not significant. Visual inspection of the plotted predicted values shows that parents who responded more defensively (-1 *SD* below Mean reflecting higher defensive responding) were closer to the perfect agreement line in predicting child self-reported symptoms of anxiety; however, this was not statistically significant and should therefore be interpreted with caution. Furthermore, the plotted effect of parent reported symptoms revealed a concave curve at the higher levels of symptom severity. This is in direct contrast to previous research that found a significant increase in congruence between parent-child informants at the higher levels of symptom severity compared with mild to moderate symptom severity (Canino et al., 1987). However, it is important to note that as the interaction in the current study was not significant, these interpretations are derived solely from visual inspection of the graph and represent a small effect size.

One possible explanation for these results may be due to the more covert internal nature of anxiety symptoms so that parents have more difficulty accurately observing and monitoring problems. Furthermore, as the age range for the study only included youth ages 11-16, this may have influenced the discrepancy in reporting. Specifically, the developmental shift during early adolescence associated with greater independence from parents may produce more discrepancy in children with internalizing problems as the combination of less frequent contact and more

70

covert symptoms interacts. Finally, it is possible that the higher levels of symptom severity themselves may account for some of the discrepancy in informant reports. While symptom validity measures were not given to the children in this sample, highly anxious children may have by the very nature of their symptoms, over-reported their anxiety during the assessment. Especially for socially anxious youth, a lengthy testing session within a novel location may have inflated their self-report ratings compared to parent ratings.

The other internalizing construct domain, depression symptoms, produced different results as analyses indicated that only the simple linear effect of parent rated depression symptoms significantly predicted child rated self-depression symptoms in the polynomial regression model. Therefore, parent's ratings of their child's depression symptoms reflected near perfect linear agreement across levels of symptom endorsement between informants that was not associated with parental defensive responding. Higher levels of parental defensive responding (-1 *SD* below Mean reflecting higher defensive responding) were expected to predict worse agreement between parent-child informants; however, results provided no evidence to support this hypothesis. Similar to the rationale provided for anxiety symptoms, adolescent youth may attempt to distance themselves from their parents as they rely more heavily on peers for emotional support. Depression symptoms may also present as more covert even when compared to anxious youth. Furthermore, societal beliefs about "moody" teenagers may influence parents in their expectations that their child's symptoms are not attributable to depression but rather reflect typical development. Therefore, these parents may be less likely to endorse symptoms of depression on a self-report measure. Semi-structured diagnostic interviews may be able to reduce this behavior with follow-up questions or clarification.

Results also differed between the externalizing symptom models. Results from the inattention polynomial regression model revealed that parent reported child inattention symptoms and child self-reported inattention symptoms were highly congruent when parents responded less defensively (+1 *SD* above Mean reflecting lower defensive responding). Visual inspection of the plotted predicted values shows that parent-child agreement for inattention symptoms was lower when parents responded more defensively (-1 *SD* below Mean reflecting higher defensive responding); however, as there was no significant interaction, these results should be interpreted with caution. Specifically, higher defensive responding (-1 *SD* below Mean reflecting higher defensive responding) parents reported fewer child inattention symptoms compared to child-self reported inattention symptoms across the lower values of child inattention; however, at higher symptom values this pattern was reversed so that parents reported greater child inattention symptoms compared to child-self reported inattention symptoms. This suggests that higher parental defensive responding (-1 *SD* below Mean reflecting higher defensive responding) only resulted in under-reporting symptoms compared to their child's self-ratings up to a certain level of inattention symptoms. This may reflect the more externalizing behaviors associated with the construct, where parents are able to minimize their child's symptoms but only until the overt nature of impairment is no longer concealable.

Furthermore, as externalizing symptoms are more likely to be observed by other individuals, the desire to minimize difficulties may reflect a pattern similar to that found in the studies conducted with mothers whose children were enrolled in a weight management program. Specifically, the mothers reported that they were acutely aware that their children were overweight; however, they expressed feeling the need to minimize their child's weight problem to protect them from the negative societal judgment and blame that they themselves had also

experienced as being made to feel like a bad mother (Jackson et al., 2007). Similarly, parents who may have historically experienced negative societal reactions as a direct result of their child's behavioral symptoms (e.g., expressing hyperactive or impulsive symptoms in the grocery store) may attempt to minimize their impairment. In addition, as the inattention symptom domain is correlated ($r = .80$; Achenbach & Rescorla, 2001) with the ADHD diagnostic category, this construct represents one of the few areas where external incentives (e.g., school accommodations such as extended time on tests and assignments, access to stimulant medication) can be obtained for the presence of identified concerns (Rabiner, 2013). Therefore, a parents desire to minimize their child's behavioral difficulties may reach a threshold at which their desire to protect the child and themselves from interpersonal judgment is mitigated by the available resources.

The aggression model resulted in a significant interaction effect represented by parent reported symptom severity moderating the quadratic form of parental defensive responding. There was also an additional quadratic effect for parent reported aggression symptoms; however, inspection of the plotted predicted values showed that the interaction effect better described the overall data model so these findings were not plotted separately. The plotted predicted values showed that parent's ratings were more discrepant at both the higher and lower levels of symptom severity and only around average values of reported symptoms did parent-child reports exhibit stronger congruence. The findings from this model, perhaps more so than from any other, support the recent push to reject the use of standardized difference scores for analyzing informant discrepancies. The plotted predicted values from the analyses provided a significantly greater amount of information for understanding the differences across raters. While these findings remain exploratory, they may allow for greater theoretical understanding of the underlying relationships between discrepancy models.

Whereas initial hypotheses predicted an interaction between defensive responding and parent reports of their child's symptoms, the aggression domain represents perhaps the most overt rationale for this behavior. Parents may experience significant societal stigma over having the "bad kid" and therefore attempt to minimize their symptom reports. Furthermore, the recent study conducted by Laird and De Los Reyes showed that when both parents and children agreed about the presence of aggressive symptoms, the children exhibited worse outcomes (2013). Finally, these results are supported by recent meta-analytic data that indicated higher rates of agreement for externalizing behaviors with more highly observable behavioral characteristics and symptoms (De Los Reyes & Kazdin, 2015).

Strengths

Results indicated that an embedded measure of validity testing was able to detect discrepancies in agreement between parent-child informants. Furthermore, analyzing the interaction terms within a polynomial regression model revealed novel information about how parental desirable responding may influence parent ratings across different symptom domains. One of the greatest problems with a lack of agreement between parent-child ratings is the possibility that under endorsement by parents will result in denied services for their child who may truly need them. As children, especially at a younger age, are not capable of seeking services themselves, parents who view symptoms as less critical than their children may be less likely to utilize services or may be less responsive to their needs.

Research has shown that low levels of parental responsiveness is associated with poorer lifetime outcomes in children (Baumrind, 1991). Therefore, high ratings of parental defensive responding (e.g. higher defensive responding was reflected in the current study as -1 *SD* below the Mean) may also represent an underlying difficulty with understanding and meeting the

psychological needs of their child. By utilizing a brief assessment measure that is already commonly used across a number of clinical settings, this study also provided evidence for its practical utility.

Limitations

Several limitations should be noted for the current study. First, one potential predictor for a parent's willingness to respond accurately may be in the reason for their initial referral. Parents seeking accommodations for their children in school may be more likely to respond openly regarding symptom presentations. Conversely, those parents who were told to seek services by a teacher or school representative may feel forced into attending an assessment appointment and may respond in a biased manner. As this factor was not tested in the current study, it remains a potential confounding variable for parental defensive responding.

Similarly, the particular sample of assessment and treatment seeking parents included in this study represents an overall restriction that may not generalize to other populations. Specifically, by utilizing an outpatient training facility, more complex cases or significantly impaired children were excluded by nature of the clinic environment. Therefore, future studies should analyze informant discrepancies across more diverse populations. Finally, the study tested a broad array of psychological difficulties but did not assess more in depth factors associated with any one domain. Therefore, more accurate information may be gleaned from focusing on one diagnostic area and potentially including more factors that could influence parental responses (e.g., parental psychopathology, age, gender, reason for referral) across a specific symptom construct.

Conclusions and Future Research

The novelty of the current findings highlight the importance of focusing on identifying alternative areas where discrepancies may exist rather than applying new techniques to old ideas. Specific attention should be given to validity measures that could detect meaningful and significant confounds within the child assessment process. These measures are likely to be more accepted if applied to parents as raters rather than the child themselves, given the persistence of the belief that children do not engage in invalidating behaviors during the assessment process.

References

Abidin, R. R. (1990). *Parenting Stress Index*. Charlottesville, VA: Pediatric Psychology Press.

Abidin, R. R. (1995). *Parenting Stress Index manual* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

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.

Ackerman, J. J., & Pritzl, T. B. (2011). Child custody evaluation practices: A 20-year follow-up. *Family Court Review, 49*, 618-628.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Alexander, F., & Selesnick, S. (1966). *The history of psychiatry*. New York, NY: New American.

American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: Author.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Andrews, V. C., Garrison, C. Z., Jackson, K. L., Addy, C. L., & McKeown, R. E. (1993). Mother-adolescent agreement on the symptoms and diagnoses of adolescent depression and conduct disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *32*(4), 731-738.

Angermeyer, M. C., Beck, M., Dietrich, S., & Holzinger, A. (2004). The stigma of mental illness: Patients' anticipations and experiences. *International Journal of Social Psychiatry, 50*, 153-162.

Angold, A., Weissman, M. M., John, K., Merikancas, K. R., Prusoff, B. A., Wickramaratne, P., ... & Warner, V. (1987). Parent and child reports of depressive symptoms in children at low and high risk of depression. *Journal of Child Psychology & Psychiatry*, *28*(6), 901-915.

Arce, R., Farina, F., Seijo, D., & Novo, M. (2015). Assessing impression management with the MMPI-2 in child custody litigation. *Assessment, 22*(6), 769-777.

Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence, 11*(1), 56-95.

Bellanca, F. F., & Pote, H. (2013). Children's attitudes towards ADHD, depression, and learning disabilities. *Journal of Research in Special Educational Needs, 13*, 234–241.

Bernreuter, R. G. (1940). The present status of personality trait tests. *Educational Record, 21*, 160-171.

Berry, W. D. (1993). *Understanding regression assumptions*. Newbury Park, CA: Sage.

Bird, H. R., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*, 78-85.

Boone, K. B. (2007). *Assessment of feigned cognitive impairment: A neuropsychological perspective*. New York, NY: Guilford Press.

Boyle, M. H., & Pickles, A. (1997). Maternal depressive symptoms and ratings of emotional disorder symptoms in children and adolescents. *Journal of Child Psychology & Psychiatry*, *38*(8), 981-992.

Bradbury, T. N., & Fincham, F. D. (1990). Attributions in marriage: Review and critique. *Psychological Bulletin*, *107*(1), 3-33.

Bray, J. H., Enright, M. F., & Easling, I. (2004). Psychological practice in rural primary care. In R. G. Frank, S. H. McDaniel, J. H. Bray, & M. Heldring (Eds.), *Primary care psychology* (pp. 243-258). Washington, DC: American Psychological Association.

Breton, J., Bergeron, L., Valla, J., Lepine, S., Houde, L., & Gaudet, N. (1995). Do children aged 9 through 11 years understand the DISC version 2.25 questions? *Journal of the American Academy of Child & Adolescent Psychiatry, 34*, 946-956.

Briggs-Gowan, M. J., Carter, A. S., & Schwab-Stone, M. (1996). Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology*, *24*(6), 749-765.

Brown, B. B. (1990). Peer groups and peer cultures. In S. S. Feldman & G. R. Elliot (Eds.), *At the threshold: The developing adolescent* (pp. 171-196). Cambridge, MA: Harvard University.

Budd, K. S. (2001). Assessing parenting competence in child protection cases: A clinical practice model. *Clinical Child & Family Psychology Review, 1*, 1-18.

Bugental, D. B. (2004). Thriving in the face of early adversity. *Journal of Social Issues*, *60*(1), 219-235.

Campis, L. K., Lyman, R. D., & Prentice-Dunn, S. (1986). The parental locus of control scale: Development and validation. *Journal of Clinical Child Psychology*, *15*(3), 260-267.

Canino, G. J., Bird, H. R., Rubio-Stipec, M., Woodbury, M. A., Ribera, J. C., Huertas, S. E., & Sesman, M. J. (1987). Reliability of child diagnosis in a Hispanic sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 26*(4), 560-565.

Cantwell, D. P. (1996). Classification of child and adolescent psychopathology. *Journal of Child Psychology & Psychiatry, 37*, 3-12.

Cantwell, D. P., Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1997). Correspondence between adolescent report and parent report of psychiatric diagnostic data. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 610-619.

Carlston, D., & Ogles, B. (2006). The impact of items and anchors on parent–child reports of problem behavior. *Child & Adolescent Social Work Journal*, *23*(1), 24-37.

Carone, D. A., & Bush, S. S. (2013). *Mild traumatic brain injury: Symptom validity assessment and malingering.* New York, NY: Springer.

Carr, G. D., Moretti, M. M., & Cue, B. J. H. (2005). Evaluating parenting capacity: Validity problems with the MMPI-2, PAI, CAPI, and ratings of child adjustment. *Professional Psychology: Research & Practice, 36*(2), 188-196.

Castaldi, J. (1988). *The relationship of maternal defensiveness to reported levels of parenting stress* (Doctoral dissertation).

Chi, T. C., & Hinshaw, S. P. (2002). Mother–child relationships of children with ADHD: The role of maternal depressive symptoms and depression-related distortions. *Journal of Abnormal Child Psychology, 30*, 387–400.

Choudhury, M. S., Pimentel, S. S., & Kendall, P. C. (2003). Childhood anxiety disorders: Parent–child (dis) agreement using a structured interview for the DSM-IV. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(8), 957-964.

Cloniger, C. R. (1999). A new conceptual paradigm from genetics and psychobiology for the science of mental health. *Australian & New Zealand Journal of Psychiatry, 33*, 174-186.

Cockeram, W. (1981). *Sociology of mental disorder*. Englewood Cliffs, NJ: Prentice-Hall.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.

Coleman, D., Walker, J. S., Lee, J., Friesen, B. J., & Squire, P. N. (2009). Children's beliefs about causes of childhood depression and ADHD: A study of stigmatization. *Psychiatric Services*, *60*, 950–957.

Collishaw, S., Goodman, R., Ford, T., Rabe-Hesketh, S., & Pickles, A. (2009). How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *Journal of Child Psychology & Psychiatry*, *50*(5), 571-580.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York, NY: Chapman & Hall.

Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist, 59*(7), 614-625.

Corrigan, P. W., & Miller, F. E. (2004). Shame, blame, and contamination: A review of the impact of mental illness stigma on family members. *Journal of Mental Health, 13*, 537-548.

Cortina, J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. *Journal of Management, 19*(4), 915-922.

Crocker, J., Quinn, D. M., & Steele, C. (1998). Social stigma. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology: Vol. 2* (4th ed., pp. 504-553). Boston, MA: McGraw-Hill.

Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine, 78*(8), 775-780.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business Psychology*, *29*, 1-19.

De la Osa, N., Ezpeleta, L., Doménech, J. M., Navarro, J. B., & Losilla, J. M. (1997). Convergent and discriminant validity of the structured diagnostic interview for children and adolescents (DICA-R). *Psychology in Spain*, *1*, 37-44.

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*(4), 858-900.

De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, *16*(3), 330-334.

DeRight, J., & Carone, D. A. (2013). Assessment of effort in children: A systematic review. *Child Neuropsychology, 21*(1), 1-24.

Derogatis, L. R. (1992). *SCL-90-R: Administration, scoring & procedures manual-II for the R (Revised) version and other instruments of the psychopathology rating scale series*. Minneapolis, MN: Clinical Psychometric Research.

Dixon, C., Murray, C., & Daiches, A. (2013). A qualitative exploration into young children's perspectives and understandings of emotional difficulties in other children. *Clinical Child Psychology & Psychiatry, 18*, 72–90.

Donohue, B., Holland, J. M., Lopez, K., Urgelles, J., & Allen, D. N. (2014). Examination of illicit drug use frequency using multiple drug assessment methods in mothers referred to treatment by child protective services. *Journal of Family Violence*, *29*(8), 911-919.

Downey, R. G., & King, C. V. (1998). Missing data in likert scales: Comparison of replacement methods. *Journal of General Psychology, 125*, 175-191.

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science & Practice, 7*, 435-453.

Durbin, C. E., & Wilson, S. (2012). Convergent validity of and bias in maternal reports of child emotion. *Psychological Assessment*, *24*(3), 647-660.

Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development, 56*, 265-275.

Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and a proposed alternative. *Organizational Behavior & Human Decision Processes, 58*, 51-100.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*, 264-286.

Engel, N. A., Rodrigue, J. R., & Geffken, G. R. (1994). Parent-child agreement on ratings of anxiety in children. *Psychological Reports*, *75*(3), 1251-1260.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using
    G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods,
    41*(4), 1149-1160.

Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2004). Parent-adolescent disagreement
    regarding psychopathology in adolescents from the general population as a risk factor for
    adverse outcome. *Journal of Abnormal Psychology, 113*, 198-206.

Fergusson, D. M., Lynskey, M. T., & Horwood, L. J. (1993). The effect of maternal depression
    on maternal ratings of child behavior. *Journal of Abnormal Child Psychology, 21*, 245-
    269.

Fitzmaurice, G. M., Laird, N. M., Zahner, G. E., & Daskalakis, C. (1995). Bivariate logistic
    regression analysis of childhood psychopathology ratings using multiple
    informants. *American Journal of Epidemiology*, *142*(11), 1194-1203.

Foley, D. L., Rutter, M., Angold, A., & Pickles, A. (2005). Making sense of informant
    disagreement for overanxious disorder. *Journal of Anxiety Disorders, 19*, 193-210.

Foucault, M. (1965). *Madness and civilization: A history of insanity in the age of reason*. New
    York, NY: Random House.

Ganzach, Y. (1997). Configurality in judgment: Is it a bias? *Psychonomic Bulletin & Review,
    4*(3), 382-386.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*.
    Washington, DC: American Psychological Association.

Goffman, E. (1963). *Stigma, notes on the management of spoiled identity*. New York, NY:
    Simon & Schuster.

Gray, D. E. (1993). Perceptions of stigma: The parents of autistic children. *Sociology of Health
    & Illness, 15*, 102-120.

Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship
    research: A conceptual primer. *Personal Relationships, 6*, 505-518.

Grills, A. E., & Ollendick, T. H. (2002). Issues in parent-child agreement: The case of structured
    diagnostic interviews. *Clinical Child & Family Psychology Review, 5*, 57-83.

Grills, A. E., & Ollendick, T. H. (2003). Multiple informant agreement and the anxiety disorders
    interview schedule for parents and children. *Journal of the American Academy of Child &
    Adolescent Psychiatry 42*(1), 30-40.

Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, NJ: Wiley.

Handwerk, M. L., Larzelere, R. E., Soper, S. H., & Friman, P. C. (1999). Parent and child discrepancies in reporting severity of problem behaviors in three out-of-home settings. *Psychological Assessment, 11*, 14–23.

Harrison, A. G., Rosenblum, Y., & Currie, S. (2010). Examining unusual digit span performance in a population of postsecondary students assessed for academic difficulties. *Assessment, 17*, 283-293.

Hartshorne, H., & May, M. A. (1928). *Studies in deceit*. New York, NY: Macmillan.

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

Hathaway, S. R., & McKinley, J. C. (1942). A Multiphasic personality schedule (Minnesota): III. the measurement of symptomatic depression. *Journal of Psychology, 14,* 73-84.

Hathaway, S. R., & McKinley, J. C. (1943a). *Manual for the Minnesota multiphasic personality inventory*. New York, NY: The Psychological Corporation.

Hathaway, S. R., & McKinley, J. C. (1943b). *The Minnesota multiphasic personality inventory* (rev. ed.). Minneapolis, MN: University of Minnesota.

Hathaway, S. R., McKinley, J. C., & MMPI Restandardization Committee. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory-2: Manual for administration and scoring*. University of Minnesota Press.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23*, 1093-1129.

Herjanic, B., Herjanic, M., Brown, F., & Wheatt, T. (1975). Are children reliable reporters? *Journal of Abnormal Child Psychology, 3*, 41-48.

Herjanic, B., & Reich, W. (1982). Development of a structured psychiatric interview for children: Agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychology, 10*, 307-324.

Hinshaw, S. P. (2007). *The mark of shame: Stigma of mental illness and an agenda for change.* Oxford, United Kingdom: Oxford University Press.

Hodges, K., & Cools, J. N. (1990). Structured diagnostic interviews. In A. M. La Greca (Ed.), *Through the eyes of the child: Obtaining self-reports from children and adolescents* (pp. 109-149). Boston, MA: Allyn and Bacon.

Hodges, K., McKnew, D., Cytryn, L., Stern, L., & Kline, J. (1982). The Child Assessment Schedule: Parent-child agreement and validity measures. *Journal of Child Psychology &. Psychiatry*, *28*, 455-466.

Holmbeck, G. N., Li, S. T., Schurman, J. V., Friedman, D., & Coakley, R. M. (2002). Collecting and managing multisource and multimethod data in studies of pediatric populations. *Journal of Pediatric Psychology*, *27*(1), 5-18.

Human, L. J., Dirks, M. A., DeLongis, A., & Chen, E. (2016). Congruence and incongruence in adolescents' and parents' perceptions of the family: Using response surface analysis to examine links with adolescents' psychological adjustment. *Journal of Youth & Adolescence, 45*, 2022-2035.

Humm, D. G., & Wadsworth, G. W. (1935). The Humm-Wadsworth Temperament Scale. *American Journal of Psychiatry, 92*, 163-200.

Jackson, D., Wilkes, L., & McDonald, G. (2007). 'If I was in my daughter's body I'd be feeling devastated': Women's experiences of mothering an overweight or obese child. *Journal of Child Health Care, 11*, 29-39.

Jameson, J. P., & Blank, M. B. (2007). The role of clinical psychology in rural mental health services: Defining problems and developing solutions. *Clinical Psychology: Science & Practice, 14*, 283-298.

Jensen, P. S., Rubio-Stipec, M., Canino, G., Bird, H. R., Dulcan, M. K., Schwab-Stone, M. E., et al. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry, 38*, 1569-1579.

Jensen, P. S., Xenakis, S. N., Davis, H., & Degroot, J. (1988). Child psychopathology rating scales and interrater agreement: Child and family characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry, 27*, 451-461.

Jensen-Doss, A., Youngstrom, E. A., Youngstrom, J. K., Feeny, N. C., & Findling, R. L. (2014). Predictors and moderators of agreement between clinical and research diagnoses for children and adolescents. *Journal of Consulting & Clinical Psychology*, *82*(6), 1151-1162.

Kasius, M. C., Ferdinand, R. F., van den Berg, H., & Verhulst, F. C. (1997). Associations between different diagnostic approaches for child and adolescent psychopathology. *Journal of Child Psychology & Psychiatry, 38*, 625-632.

Kaushik, A., Kostaki, E., & Kyriakopoulos, M. (2016). The stigma of mental illness in children and adolescents: A systematic review. *Psychiatry Research, 243*, 469-494.

Kazdin, A. E. (1989). Developmental psychopathology: Current research, issues, and directions. *American Psychologist, 44*, 180-187.

Kelley, M. L., Bravo, A. J., Hamrick, H. C., Braitman, A. L., White, T. D., & Jenkins, J. (2017). Parents' reports of children's internalizing symptoms: Associations with parents' mental health symptoms and substance use disorder. *Journal of Child & Family Studies, 26,* 1646-1654.

Kirkwood, M. W. (2012). Overview of tests and techniques to detect negative response bias in children. In E. M. S. Sherman & B. L. Brooks (Eds.), *Pediatric forensic neuropsychology* (pp. 136-161). New York, NY: Oxford University Press.

Kirkwood, M. W., Kirk, L. W., Blaha, R. Z., & Wilson, P. (2010). Noncredible effort during pediatric neuropsychological exam: A case series and literature review. *Child Neuropsychology, 16*, 604-618.

Klein, R. G. (1991). Parent-child agreement in clinical assessment of anxiety and other psychopathology: A review. *Journal of Anxiety Disorders, 5*, 187-198.

Klein, D. N., Dougherty, L. R., & Olino, T. M. (2005). Toward guidelines for evidence-based assessment of depression in children and adolescents. *Journal of Clinical Child & Adolescent Psychology, 34*, 412-432.

Klein, D. N., Ouimette, P. C., Kelly, H. S., Ferro, T., & Riso, L. P. (1994). Test-retest reliability of team consensus best-estimate diagnoses of Axis I and II disorders in a family study. *American Journal of Psychiatry, 151*(7), 1043-1047.

Kleinbaum, D., Kupper, L., Nizam, A., & Rosenberg, E. (2013). *Applied regression analysis and other multivariable methods*. Scarborough, Canada: Nelson Education.

Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and non-clinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology & Psychiatry*, *34*(6), 991-1006.

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry, 160*, 1566-1577.

Kroes, G., Veerman, J. W., & De Bruyn, E. E. (2003). Bias in parental reports? Maternal psychopathology and the reporting of problem behavior in clinic-referred children. *European Journal of Psychological Assessment*, *19*(3), 195-203.

Kuehnle, K., Coulter, M., & Firestone, G. (2000). Child protection evaluations: The forensic stepchild. *Family & Conciliation Courts Review, 38*, 368-391.

LaFiosca, T., & Loyd, B. H. (1986). Defensiveness and the assessment of parental stress and anxiety. *Journal of Clinical Child Psychology*, *15*(3), 254-259.

Lai, B. S., Beaulieu, B., Ogokeh, C. E., Self-Brown, S., & Kelley, M. L. (2015). Mother and child reports of hurricane related stressors: Data from a sample of families exposed to hurricane Katrina. *Child & Youth Care Forum, 44*, 549-565.

Laird, R. D., & De Los Reyes, A. (2013). Testing informant discrepancies as predictors of early adolescent psychopathology: Why difference scores cannot tell you what you want to know and how polynomial regression may. *Journal of Abnormal Child Psychology*, *41*(1), 1-14.

Laird, R. D., & LaFleur, L. K. (2016). Disclosure and monitoring as predictors of mother–adolescent agreement in reports of early adolescent rule-breaking behavior. *Journal of Clinical Child & Adolescent Psychology*, *45*(2), 188-200.

Laird, R. D., & Weems, C. F. (2011). The equivalence of regression models using difference scores and models using separate scores for each informant: Implications for the study of informant discrepancies. *Psychological Assessment, 23*(2), 388-397.

Lambert, C. E., Arbuckle, S. A., & Holden, R. R. (2016). The Marlowe–Crowne Social Desirability Scale outperforms the BIDR Impression Management Scale for identifying fakers. *Journal of Research in Personality*, *61*, 80-86.

Lapouse, R., & Monk, M. A. (1958). An epidemiologic study of behavior characteristics in children. *American Journal of Public Health, 48*, 1134-1144.

Larrabee, G. J. (2007). Introduction: Malingering, research designs and base rates. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 3-13). New York, NY: Oxford University Press.

LeBreton, J. M., Tonidandel, S., & Krasikova, D. V. (2013). Residualized relative importance analysis: A technique for the comprehensive decomposition of variance in higher order regression models. *Organizational Research Methods, 16*(3), 449-473.

Leung, J. T. Y., Shek, D. T. L., & Li, L. (2016). Mother-child discrepancy in perceived family functioning and adolescent developmental outcomes in families experiencing economic disadvantage in Hong Kong. *Journal of Youth & Adolescence, 45*, 2036-2048.

Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology, 27,* 363-385.

Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *Journal of Clinical Child Psychology, 19*, 136-143.

Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Social Psychology, 140*, 628-635.

MacCallum, F., Golombok, S., & Brinsden, P. (2007). Parenting and child development in families with a child conceived through embryo donation. *Journal of Family Psychology, 21*(2), 278-287.

Maller, J. B. (1930). The effect of signing one's name. *School & Society, 31*, 882-884.

Maller, J. B. (1932). *Character sketches*. New York, NY: Bureau of Publications, Teachers College, Columbia University.

Maller, J. B. (1944). Personality tests. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (pp. 170-213). New York, NY: Ronald Press.

Marshall, P., Schroeder, R., O'Brien, J., Fischer, R., Ries, A., Blesi, B., & Barker, J. (2010). Effectiveness of symptom validity measures in identifying cognitive and behavioral symptom exaggeration in adult attention deficit hyperactivity disorder. *The Clinical Neuropsychologist*, *24*(7), 1204-1237.

Matthews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology, 1*, 167-195.

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

Meehl, P. E. (1945). The dynamics of structured personality tests. *Journal of Clinical Psychology, 1*, 296-303.

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30*, 525-564.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128-165.

Müller, J. M., & Furniss, T. (2013). Correction of distortions in distressed mothers' ratings of their preschool children's psychopathology. *Psychiatry Research*, *210*(1), 294-301.

Müller, J. M., Romer, G., & Achtergarde, S. (2014). Correction of distortion in distressed mothers' ratings of their preschool-aged children's Internalizing and Externalizing scale score. *Psychiatry Research*, *215*(1), 170-175.

Najman, J. M., Williams, G. M., Nikles, J., Spence, S., Bor, W., & O'Callaghan, M. (2000). Mothers' mental illness and child behavior problems: Cause-effect association or observation bias? *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 592-602.

Nelemans, S. A., Branje, S. J. T., Hale, W. W., Goossens, L., Koot, H. M., Oldenhinkel, A. J., & Meeus, W. H. J. (2016). Discrepancies between perceptions of the parent-adolescent relationship and early adolescent depressive symptoms: An illustration of polynomial regression analysis. *Journal of Youth & Adolescence, 45*, 2049-2063.

O'Driscoll, C., Heary, E., Hennessy, & McKeague, L. (2012). Explicit and implicit stigma towards peers with mental health problems in childhood and adolescence. *Journal of Child Psychology & Psychiatry, 53*, 1054–1062.

Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., & Lipman, E. L. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(8), 1078-1085.

Ohan, J. L., Visser, T. A. W., Moss, R. G., & Allen, N. B. (2013). Parents' stigmatizing attitudes toward psychiatric labels for ADHD and depression. *Psychiatric Services, 64*, 1270–1273.

Olson, W. C. (1936). The waiver of signature in personal reports. *Journal of Applied Psychology, 20*, 442-450.

Otto, R. K., & Collins, R. P. (1995). Use of the MMPI-2/MMPI-A in child custody evaluations. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), *Forensic applications of the MMPI-2* (pp. 222–252). Thousand Oaks, CA: Sage.

Otto, R. K., Edens, J. F., & Barcus, E. H. (2000). The use of psychological testing in child custody evaluations. *Family & Conciliation Courts Review, 38,* 312–340.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality & Social Psychology, 46*(3), 598-609.

Pellegrino, J. F., Singh, N. N., & Carmanico, S. J. (1999). Concordance among three diagnostic procedures for identifying depression in children and adolescents with EBD. *Journal of Emotional and Behavioral Disorders, 7*, 118 –127.

Peterson, E., & Peterson, R. L. (2015). Understanding deception from a developmental perspective. In M. W. Kirkwood (Ed.), *Validity testing in child and adolescent assessment* (pp. 42-61). New York, NY: Guilford Press.

Phelan, J. C., Bromet, E. J., & Link, B. G. (1998). Psychiatric illness and family stigma. *Schizophrenia Bulletin, 24*, 115-126.

Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology, 20*, 51-63.

Pilkonis, P. A., Heape, C. L., Ruddy, J., & Serrao, P. (1991). Validity in the diagnosis of personality disorders: The use of the LEAD standard. *Psychological Assessment, 3*, 46 – 54.

Rabiner, D. L. (2013). Stimulant prescription cautions: Addressing misuse, diversion, and malingering. *Current Psychiatry Reports*, *15*(7), 1-8.

Radnitz, C., & Todd, L. E. (2016). Social desirability may explain why some caregivers of overweight children report less frequent high calorie food intake. *Eating Behaviors, 23*, 48-51.

Rapee, R. M., Barrett, P. M., Dadds, M. R., & Evans, L. (1994). Reliability of the DSM-III-R childhood anxiety disorders using structured interview: Interrater and parent-child agreement. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*(7), 984-992.

Read, J., & Harre, N. (2001). The role of biological and genetic causal beliefs in the stigmatization of "mental patients." *Journal of Mental Health, 10*, 223–235.

Renk, K., Roddenberry, A., Oliveros, A., & Sieger, K. (2007). The relationship of maternal characteristics and perceptions of children to children's emotional and behavioral problems. *Child & Family Behavior Therapy*, *29*(1), 37-57.

Renouf, A. G., & Kovacs, M. (1994). Concordance between mothers' reports and children's self-reports of depressive symptoms: A longitudinal study. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 59-66.

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*, 485-499.

Rohling, M. L. (2004). Who do they think they're kidding: Review of the use of symptom validity tests with children? *Division of Clinical Neuropsychology Newsletter*, *22*, 1-26.

Rosenzweig, S. (1934). A suggestion for making verbal personality tests more valid. *Psychological Review, 41*, 400-401.

Rutter, M., & Graham, P. (1968). The reliability and validity of the psychiatric assessment of the child: Interview with the child. *British Journal of Psychiatry, 114*, 563-579.

Salbach-Andrae, H., Klinkowski, N., Lenz, K., & Lehmkuhl, U. (2009). Agreement between youth-reported and parent-reported psychopathology in a referred sample. *European Child & Adolescent Psychiatry*, *18*(3), 136-143.

Schütz, A. (1998). Assertive, offensive, protective, and defensive styles of self-presentation: A taxonomy. *The Journal of Psychology*, *132*(6), 611-628.

Schwab-Stone, M., Fallon, T., Briggs, M., & Crowther, B. (1994). Reliability of diagnostic reporting for children aged 6-11 years: A test-retest study of the Diagnostic Interview Schedule for Children-Revised. *American Journal of Psychiatry, 151*, 1048-1054.

Seiffge-Krenke, I., & Kollmar, F. (1998). Discrepancies between mothers' and fathers' perceptions of sons' and daughters' problem behavior: A longitudinal analysis of parent-adolescent agreement on internalizing and externalizing problem behavior. *Journal of Child Psychology & Psychiatry, 39*, 687-697.

Sherman, E. (2015). Terminology and diagnostic concepts. In M. W. Kirkwood (Ed.), *Validity testing in child and adolescent assessment* (pp. 22-41). New York, NY: Guilford Press.

Sibley, M. H., Campez, M., & Raiker, J. S. (2017). Reexamining ADHD-related self-reporting problems using polynomial regression. *Assessment, 24*(3), 290-296.

Silverman, W. K., & Eisen, A. R. (1992). Age differences in the reliability of parent and child reports of child anxious symptomatology using a structured interview. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(1), 117-124.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13*, 545-561.

Snarr, J. D., Slep, A. M. S., & Grande, V. P. (2009). Validation of a new self-report measure of parental attributions. *Psychological Assessment*, *21*(3), 390.

Sobol, M. P., Ashbourne, D. T., Earn, B. M., & Cunningham, C. E. (1989). Parents' attributions for achieving compliance from attention-deficit-disordered children. *Journal of Abnormal Child Psychology*, *17*(3), 359-369.

Sourander, A., Helstela, L., & Helenius, H. (1999). Parent-adolescent agreement on emotional and behavioral problems. *Social Psychiatry & Psychiatric Epidemiology, 34*, 657-663.

Southwell, O., & Fox, J. R. E. (2011). Maternal perceptions of overweight and obesity in children: A grounded theory study. *British Journal of Health Psychology, 16*, 626-641.

Spencer, D. (1938). Frankness of subjects on personality measures. *Journal of Educational Psychology, 29*, 26-35.

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry, 24*, 399-411.

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology, 22*, 107-115.

Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, *14*(3), 189-207.

Sweet, J. J., & Guidotti Breting, L. M. (2013). Symptom validity test research: Status and clinical implication. *Journal of Experimental Psychopathology, 4,* 6-19.

Sylvester, C. E., Hyde, T. S., & Reichler, R. J. (1987). The diagnostic interview for children and personality inventory for children in studies of children at risk for anxiety disorders or depression. *Journal of the American Academy of Child & Adolescent Psychiatry, 26*, 668-675.

Tarullo, L. B., Richardson, D. T., Radke-Yarrow, M., & Martinez, P. E. (1995). Multiple sources in child diagnosis: Parent-child concordance in affectively ill and well families. *Journal of Clinical Child Psychology, 24*, 173-183.

Thammasitboon, S., & Cutrer, W. B. (2013). Diagnostic decision-making and strategies to improve diagnosis. *Current Problems in Pediatric and Adolescent Health Care*, *43*(9), 232-241.

Thompson Jr., R. J., Merritt, K. A., Keith, B. R., Bennett Murphy, L., & Johndrow, D. A. (1993). The role of maternal stress and family functioning in maternal distress and mother-reported and child-reported psychological adjustment of nonreferred children. *Journal of Clinical Child Psychology, 22*, 78-84.

Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent–child relationships? *Journal of Abnormal Child Psychology, 31*, 13–27.

Van de Mortel, T. (2008). Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, 25*(4), 40-48.

Verhulst, F. C., & Van der Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology & Psychiatry, 33*, 1011-1023.

Wahl, O. F. (1999). *Telling is risky business: Mental health consumers confront stigma*. New Brunswick, N.J.: Rutgers University Press.

Walker, J. S. (2011). Malingering in children: Fibs and faking. *Child & Adolescent Psychiatric Clinics in North America*, *20*, 547-556.

Walker, J. S., Coleman, D., Lee, J., Squire, P. N., & Friesen, B. J. (2008). Children's stigmatization of childhood depression and ADHD: Magnitude and demographic variation in a national sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*, 912–920.

Weiner, B. (1995). Attribution theory in organizational behavior: A relationship of mutual benefit. *Attribution theory: An organizational perspective*, 3-6.

Weiss, M. E. (1994). Children's attitudes toward the mentally ill: An eight-year longitudinal follow-up. *Psychological Reports, 74*(1), 51-56.

Widiger, T. A., & Clark, L. A. (2000). Toward DSM-V and the classification of psychopathology. *Psychological Bulletin, 126*, 946-963.

Williams, S., McGee, R., Anderson, J., & Silva, P. A. (1989). The structure and correlates of self-reported symptoms in 11-year old children. *Journal of Abnormal Child Psychology, 17*, 55-71.

Williams, S. L., & Polaha, J. (2014). Rural parents' perceived stigma of seeking mental health services for their children: Development and evaluation of a new instrument. *Psychological Assessment, 26*(3), 763-773.

Wolraich, M. L., Lambert, E. W., Bickman, L., Simmons, T., Doffing, M. A., & Worley, K. A. (2004). Assessing the impact of parent and teacher agreement on diagnosing attention-deficit hyperactivity disorder. *Journal of Developmental & Behavioral Pediatrics, 25*, 41-47.

Yaban, E. H., Sayil, M., & Tepe, Y. K. (2014). Are discrepancies in perceptions of psychological control related to maladjustment? A study of adolescents and their parents in Turkey. *International Journal of Behavioral Development, 38*(6), 550-562.

Yeh, M., & Weisz, J. R. (2001). Why are we here at the clinic? Parent-child (dis)agreement on referral problems at outpatient treatment entry. *Journal of Consulting & Clinical Psychology, 69*(6), 1018-1085.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology & Psychology, 18*, 459-482.

Zilboorg, G., & Henry, G. W. (1941). *A history of medical psychology*. New York, NY: Norton.

Vita

Megan Lilly is from all over the southern United States. She earned a B.S. in Psychology from The University of New Orleans in 2009. After two years serving as a research assistant at Tulane University she enrolled as a graduate student at Louisiana State University where she earned her M.A. in Clinical Psychology in 2015. Megan completed her predoctoral internship at Duke University School of Medicine in August 2017 with a concentration in pediatric psychology.