

# Integrating Multiple Informants' Reports: How Conceptual and Measurement Models May Address Long-Standing Problems in Clinical Decision-Making

Clinical Psychological Science  
2020, Vol. 8(6) 953–970  
© The Author(s) 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/2167702620924439  
www.psychologicalscience.org/CPS  


Bridget A. Makol<sup>1</sup>, Eric A. Youngstrom<sup>2</sup>, Sarah J. Racz<sup>1</sup>,  
Noor Qasmieh<sup>1</sup>, Lara E. Glenn<sup>3</sup>, and Andres De Los Reyes<sup>1</sup>

<sup>1</sup>Department of Psychology, University of Maryland at College Park; <sup>2</sup>Department of Psychology, University of North Carolina at Chapel Hill; and <sup>3</sup>Department of Psychology, University of Wyoming

## Abstract

Assessing youth psychopathology involves collecting multiple informants' reports. Yet multi-informant reports often disagree, which necessitates integrative strategies that optimize predictive power. The trait-score approach leverages principal components analysis to account for the context and perspective from which informants provide reports. This approach may boost the predictive power of multi-informant reports and thus warrants rigorous testing. We tested the trait score approach using multi-informant reports of adolescent social anxiety in a mixed clinical and community sample of adolescents ( $N = 127$ ). The trait score incrementally predicted observed social anxiety ( $\beta_s = 0.47\text{--}0.67$ ) and referral status (odds ratios = 2.66–6.53) above and beyond individual informants' reports and a composite of informants' reports. The trait score predicted observed behavior at magnitudes well above those typically observed for individual informants' reports of internalizing psychopathology (i.e.,  $r_s = .01\text{--}.15$ ). Findings demonstrate the ability of the trait score to improve prediction of clinical indices and potentially transform widely used practices in multi-informant assessments.

## Keywords

multiple informants, assessment, integrative strategies, principal components analysis, social anxiety

Received 11/7/19; Revision accepted 4/3/20

Imagine you are a clinician working with an adolescent who reports a consistent and impairing fear of social situations, including an intense fear of negative evaluation from others. Behaviorally, the adolescent avoids social situations and in particular those that involve interacting with unfamiliar individuals. When unable to avoid these situations, the adolescent reports enduring them with great distress. In contrast, the adolescent's parent reports minimal levels of social anxiety and instead identifies behaviors displayed at home that result in the adolescent underperforming at school (e.g., poor schoolwork, rule breaking) as important to focus on in treatment. As a clinician, how will you use this information to inform the care that you provide? How will you conceptualize the adolescent's social anxiety, and what diagnoses will you assign? Is one informant "right" and the other "wrong"? As you begin

treatment planning, which adolescent problems ought to be targeted first?

## Suboptimal Use of Multi-Informant Reports: A Clinical Decision-Making Problem

In this article, we address commonly encountered questions in clinical practice by advancing an innovative strategy for integrating multi-informant reports of youths' mental

### Corresponding Author:

Andres De Los Reyes, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland at College Park, Biology/Psychology Building, Room 3123H, College Park, MD 20742  
E-mail: adlr@umd.edu

health. Indeed, we opened this article with a case example that highlights a common clinical observation: Youths vary considerably in the contexts in which they display mental health symptoms (e.g., home, school; Achenbach, McConaughy, & Howell, 1987), and significant others in their lives (e.g., parents, teachers) vary in their capacities for observing them within and across these contexts (Dirks, De Los Reyes, Briggs-Gowan, Cella, & Wakschlag, 2012). To account for these complexities in clinical presentation, clinicians rely on multiple informants to provide reports of youths' mental health (Hunsley & Mash, 2007). Yet when implementing this *multi-informant approach*, informants' reports display overall low to moderate levels of convergence (i.e.,  $r_s = .20-.30$ ; Achenbach et al., 1987; De Los Reyes et al., 2015). Consequently, how a mental health professional interprets these *informant discrepancies* has profound implications for crucial clinical decision-making tasks, including estimating prognosis, planning treatment, and evaluating treatment effectiveness (Hawley & Weisz, 2003). Yet despite decades of research attention on these topics, there is little evidence that supports the *integrated* use of multi-informant reports (De Los Reyes & Langer, 2018; Hunsley & Mash, 2007). Does integrating multi-informant data improve the accuracy, outcome, or efficiency of clinical work relative to alternative practices such as choosing a single, "optimal" informant or examining each informant's report individually? Without an answer to this question, the field lacks clear guidelines for clinicians and researchers seeking to leverage multi-informant reports to improve clinical decision-making (Beidas et al., 2015; Marsh, De Los Reyes, & Lilienfeld, 2018).

In practice, low correspondence often results in assessors shunning multi-informant approaches to assessment and instead making clinical decisions that rely on data from a single source (Loeber, Green, & Lahey, 1990; Loeber, Green, Lahey, & Stouthamer-Loeber, 1989; Marsh, Zeveney, & De Los Reyes, 2020). This practice likely reflects long-held assumptions that low informant correspondence signals measurement error or informant biases (De Los Reyes, 2011). Needless to say, the evidence suggests that no single informant's report robustly predicts clinical indices germane to service delivery. For example, any one informant's report relates at only small to moderate levels with observed youth behavior, including for externalizing problems (e.g.,  $r_s = .02-.52$ ; Becker, Luebke, Fite, Stoppelbein, & Greening, 2014; Doctoroff & Arnold, 2004; Henry & Metropolitan Area Child Study Research Group, 2006; Hinshaw, Han, Erhardt, & Huber, 1992; Wakschlag et al., 2008; Winsler & Wallace, 2002) and internalizing problems (e.g.,  $r_s = .01-.33$ ; Becker et al., 2014; Beidel,

Turner, Hamlin, & Morris, 2000; Cartwright-Hatton, Hodges, & Porter, 2003; Hinshaw et al., 1992; Winsler & Wallace, 2002). In fact, correlations between informants' reports and youths' behavior vary widely; most correlations fall in the small range (i.e., majority of  $r_s = .01-.15$ ). Crucially, practitioners and researchers in youth mental health lack instruments that consistently—and at large magnitudes—predict key clinical indices relevant to youth mental health service delivery and research. We need efficient yet comprehensive assessments that when integrated, improve decision-making in clinical and research settings (Beidas et al., 2015).

### Addressing the Clinical Decision-Making Problem: The Trait-Score Approach

Nearly 20 years ago, Kraemer and colleagues (2003) proposed an integrative approach with potential for addressing these long-standing assessment issues. We refer to this strategy as the *trait-score approach*, and it reflects an insightful observation by Kraemer and colleagues: To arrive at a clinically useful set of multi-informant reports, one must select informants who systematically vary in the contexts and perspectives from which they observe youths. Stated differently, assessors should select informants who, for systematic reasons, tend to provide discrepant reports. This reflects an innovative departure from historical thinking on low convergence in multi-informant assessments. Rather than considering low informant convergence as a *problem* for decision-making, assessors might leverage low convergence to *enhance* decision-making.

The underlying assumptions of the trait-score approach fundamentally differ from assumptions underlying alternative strategies for integrating continuous, multivariate data. For instance, researchers often take a *composite-score* approach to integrate multi-informant reports, essentially by taking the simple average among continuous scores taken from informants' reports (De Los Reyes, Alfano, Lau, Augenstein, & Borelli, 2016). The composite-score approach has its reference to the very origins of classical test theory (Borsboom, 2005). In particular, Edgeworth (1888) asserted that variations among individual raters of the same behavior reflect "error" around a "true score" representation of the behavior being assessed. In this respect, the composite-score approach treats variations among informants' reports as error. In contrast, the trait-score approach assumes that reports vary from one another both systematically and because of measurable factors.

Kraemer and colleagues (2003) advanced a mathematical model for sources of variability that contribute to an informant's rating of youth symptoms:

Informant's Rating = Trait + Context + Perspective.

Kraemer and colleagues' model harkens back to decades of research on techniques for aggregating multivariate data (e.g., factor analysis; Tabachnick & Fidell, 2001). In fact, Kraemer and colleagues' innovation is in leveraging a data-aggregation technique that has historically been used to parsimoniously aggregate *multiple item responses* on a measure of someone's behavior to, instead, parsimoniously aggregate *multiple informants' ratings* on a measure of someone's behavior. To implement their theoretical approach, Kraemer and colleagues applied principal components analysis (PCA) to a carefully curated set of multi-informant reports. This statistical approach is consistent with 50 years of research that points to robustly low levels of correspondence among informants' reports (Achenbach et al., 1987; De Los Reyes et al., 2015). Furthermore, mental health professionals commonly solicit the reports of informants who qualitatively vary in the contexts in which they observe the behaviors about which they provide reports (e.g., parents at home vs. teachers at school) as well as the perspectives from which they provide reports (e.g., self-reports vs. observer reports; De Los Reyes, Thomas, Goodman, & Kundey, 2013). The PCA approach advanced by Kraemer and colleagues requires selecting informants who vary across these systematic factors and thus assumes that the factors explaining variability among the informants' reports entered into the model are linear combinations of orthogonal variables. This approach reflects a conservative and falsifiable strategy for integrating multi-informant data. Indeed, to the degree that the data entered into the model violate assumptions of orthogonality among informants' reports, the resulting information gleaned from the model should fail to predict clinically relevant criterion variables over and above alternative approaches.

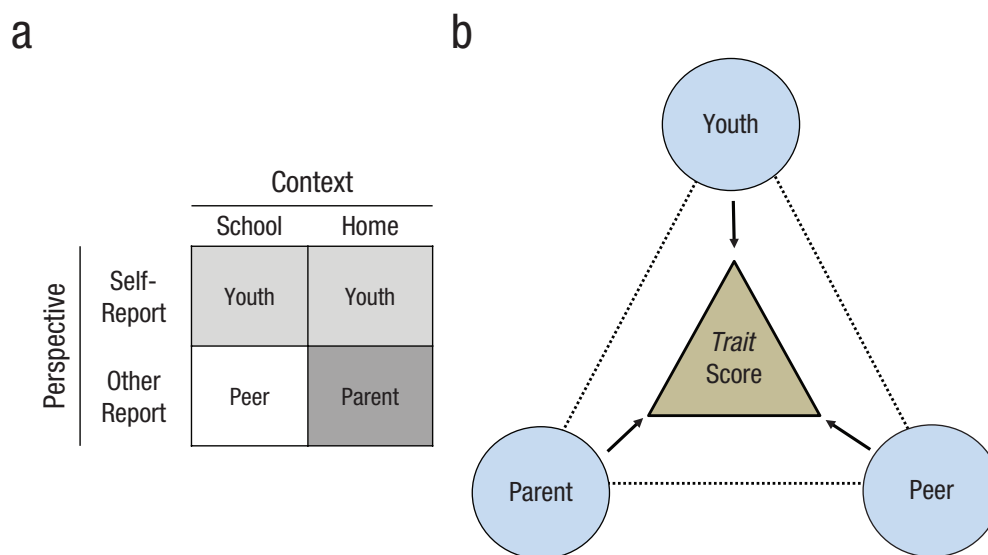
The Kraemer and colleagues (2003) approach involves synthesizing the variability among informants' reports into three sources: *context* (i.e., environment in which the informant observes youths), *perspective* (i.e., whether the report comes from an observer or self-rater), and *trait* (i.e., concerns that manifest across contexts and perspectives). This approach is designed to identify these three sources of variability among informants' reports through extraction of three components, and specific sources of variability are identified through examination of component weights. That is, Kraemer and colleagues posited that when one selects informants who vary in their contexts and perspectives, one component (i.e., trait score) reflects variability for which all three informants' reports load strongly and in the same direction. Furthermore, two additional

components reflect variability for which informants' report loadings are consistent with a context score (i.e., informants from different contexts load in opposite directions) and perspective score (i.e., self-reports load in the opposite direction of observer informants' reports). By isolating a component reflecting concerns that manifest across contexts and perspectives, Kraemer and colleagues proposed that the trait score optimizes multi-informant reports for maximal utility in predicting a wide array of clinical outcomes. This approach, which holds great promise, has yet to be rigorously tested.

### **Building an Evidence Base for Selecting Informants: The Operations Triad Model**

To optimize the predictive power of the approach, Kraemer and colleagues (2003) argued that one should use a "mix-and-match" criterion to identify those informants who systematically vary in sources of variability (Fig. 1). This approach is similar to global positioning system (GPS) technologies that use data from three satellites to triangulate on a precise location in space. One cannot achieve triangulation with data from a single GPS satellite or arrive at triangulated data using three GPS satellites situated in the same location in space. Likewise, no single "optimal informant" exists to assess youth's mental health, and three similar informants (e.g., three teachers) would not provide effective triangulation given their position in the same space or environment (e.g., school). Note that the process by which one identifies triangulated satellites requires data to ensure that the satellites are in fact situated in optimal locations in space. Likewise, the trait-score approach holds value insofar as an evidence base exists to support that low correspondence among reports reflects systematic processes. The operations triad model (OTM; De Los Reyes et al., 2013) provides a guide for building such an evidence base for multi-informant assessments. In particular, the OTM facilitates detecting circumstances in which low convergence among multi-informant reports reflects meaningful, systematic information about clinical phenomena. Thus, the OTM facilitates determining which informants might be most appropriate to use when applying the trait-score approach.

Several studies informed by the OTM reveal that systematic patterns of multi-informant reports facilitate characterizing clinical presentations. For example, relative to patterns of parent-teacher reports that diverge in estimates of youths' autism spectrum disorder symptoms, patterns of parent-teacher reports that converge on moderate or high levels of such symptoms point to



**Fig. 1.** Example of use of “mix-and-match” criterion to identify optimal informants to include in a multi-informant assessment (a). Informants systematically vary in the perspective and context from which they rate youths’ mental health symptoms, with the goal of effectively triangulating on a trait score. (b) Graphical depiction of multi-informant reports triangulating, much like the global positioning system, to identify the trait score. Both peer and parent reports provide information from an “other” perspective; peers provide information about the school context, and parents provide information about the home context. Youth reports provide the self-perspective and information about both the school and home contexts. Panel (a) is adapted from Kraemer et al. (2003).

increased clinical severity of the youths being evaluated (e.g., medication, diagnosis; Lerner, De Los Reyes, Drabick, Gerber, & Gadow, 2017). These patterns in multi-informant reports also have predictive utility: In a recent study of psychiatric inpatient intake assessments, patterns of agreement or disagreement in parent-adolescent reports of adolescent affective symptoms facilitated prediction of treatment outcomes (Makol, De Los Reyes, Ostrander, & Reynolds, 2019). In fact, an emerging body of evidence indicates that informants who vary in their contexts and perspectives provide meaningful information about cross-contextual consistencies in behavior (De Los Reyes, Henry, Tolan, & Wakschlag, 2009; Hartley, Zakriski, & Wright, 2011; Kwon, Kim, & Sheridan, 2012), mental health service use (Jones et al., 2019; Makol & Polo, 2018), and treatment outcomes (Becker-Haimes, Jensen-Doss, Birmaher, Kendall, & Ginsburg, 2017). In sum, research across psychopathology domains, informants, measurement methods, and developmental periods support that multi-informant reporting patterns yield meaningful clinical information. This work supports a key assumption of Kraemer and colleagues’ (2003) trait-score approach: Variations among informants’ reports reflect systematic factors germane to the clinical presentations of those assessed. In turn, informants should be selected for the unique information they provide in an effort to optimize the utility of youth mental health assessments.

## The Need for Rigorous, Conceptually Grounded Research

Despite the promise of Kraemer and colleagues’ (2003) approach, it has been underused and has yet to be rigorously examined. This is due largely to the lack of conceptual frameworks and basic science to guide researchers in selecting appropriate informants to leverage in assessments. Furthermore, although Kraemer and colleagues proposed that the trait score is the primary predictive tool derived from their approach, it remains an empirical question as to whether the context and perspective scores also yield valuable clinical information. Beyond these conceptual issues, practical challenges have curtailed rigorous tests of trait scores. Given resource limitations, assessments rarely include the three informants necessary to implement the approach. This occurs despite the fact that evidence-based assessment of youths’ mental health entails collecting multi-informant reports (Hunsley & Mash, 2007). Even among studies that have estimated trait scores, use of nonparallel measures across informants hinders the interpretability of study findings (e.g., Perren, Stadelmann, Burgin, & Von Klitzing, 2006). Furthermore, prior trait-score studies have been largely limited to use of parent, youth, and teacher reports (Noordhof, Oldehinkel, Verhulst, & Ormel, 2008; Perren et al., 2006) and examinations of youths’ externalizing problems

(Houts, Caspi, Pianta, Arseneault, & Moffitt, 2010). Finally, no study to our knowledge has examined whether the trait score displays incremental validity in predicting independent criterion variables (e.g., observed behavior, clinical status) over and above alternative strategies for either examining informants' reports individually or aggregating data (e.g., composite score).

## Current Study

Using Kraemer and colleagues' (2003) approach, we surmised that a trait score derived from a set of strategically selected informants will significantly boost the power of multi-informant assessments to predict criterion variables taken using other modalities over and above the levels of correspondence typically observed among informants' reports and these modalities (i.e.,  $r_s = .01-.15$ ; Becker et al., 2014; Beidel, Turner, Hamlin, & Morris, 2000; Cartwright-Hatton et al., 2003; Hinshaw et al., 1992; Winsler & Wallace, 2002). In this study, we applied the trait-score approach to social anxiety assessments in a mixed clinical and community sample of adolescents. To provide a conservative test of this approach, we examined whether the trait score predicted independent criterion variables above and beyond well-established measures of social anxiety taken from individual informants' reports. We addressed three aims. First, we expected the social anxiety trait score to demonstrate incremental validity in predicting observed adolescent social anxiety in controlled laboratory tasks relative to well-established measures of social anxiety taken from individual informants' reports. Second, we expected the social anxiety trait score to demonstrate incremental validity in distinguishing adolescents on referral status (i.e., clinic referred vs. community control) relative to well-established measures of social anxiety taken from individual informants' reports. Third, we expected the social anxiety trait score to provide incremental prediction of criterion variables over and above an alternative strategy for integrating multi-informant reports (i.e., the composite-score approach).

## Research Demonstration: Applying the Trait-Score Approach to Adolescent Social Anxiety

In line with the example we presented at the opening of this article, we demonstrate the utility of the trait-score approach with an innovative, evidence-based approach to multi-informant assessments of adolescent social anxiety. We seek to use this demonstration to highlight how the trait-score approach addresses two key limitations in evidence-based assessment of youth

mental health problems broadly and adolescent social anxiety in particular. First, when parents and adolescents disagree, assessors often make judgments that more closely align with parent reports relative to adolescent reports (Brown-Jacobsen, Wallace, & Whiteside, 2011; Hawley & Weisz, 2003). In social anxiety assessments, parent-adolescent discrepancies often result in the interpretation that a social desirability bias leads adolescents to downplay or underreport their social anxiety symptoms (De Los Reyes et al., 2015). In contrast, by placing an emphasis on selecting informants who systematically vary in the context and perspective from which they rate youth mental health symptoms, the trait-score approach *forces* discrepancies among reports. That is, by normalizing the presence of discrepant reports, the approach provides methods for integrating these reports.

Second, parents serve as "gatekeepers" who play a crucial role in initiating mental health care, and thus, evidence-based assessments involve collecting both parent and adolescent reports, particularly for internalizing concerns (Makol & Polo, 2018). Yet prior work indicates that relying *only* on these two informants yields insufficient clinical information. First, relative to children from earlier developmental periods, parents of adolescents have reduced access to and understanding of their adolescents' activities and relationships with peers (Smetana, 2008). Adolescent social anxiety often manifests in interactions with peers, which makes the peer context a key target of treatment (Hofmann et al., 1999). In support of these ideas, parent reports of adolescent social anxiety fail to predict adolescents' self-reported experiences in interactions with unfamiliar peers (Deros et al., 2018). Parents' lack of access to the peer context likely contributes to the low correspondence often observed between parent and adolescent reports of adolescent social anxiety. In contrast, to triangulate on the trait score, Kraemer and colleagues' (2003) approach requires three carefully selected informants who vary in context and perspective.

When assessing adolescent social anxiety, which informants should be leveraged in the trait-score approach? In their self-reports, adolescents provide information from a self-perspective and across contexts. In contrast, parents provide information about mental health concerns observed in the home context but often do not directly observe adolescents in nonhome contexts. Although teachers commonly stand in as a third informant in multi-informant assessments of youths' externalizing problems, they may not be the key third informant to leverage when applying the trait-score approach to adolescent social anxiety. Given that information about the unfamiliar peer context is crucially

lacking in assessments of adolescent social anxiety, an optimal informant would have opportunities to observe the adolescent within social interactions with unfamiliar peers. Commonly used methods for collecting peers' reports, such as peer nominations, provide information that is predictive of youths' social anxiety (Gazelle, Workman, & Allan, 2010). However, these approaches often require large groups of peers to provide valid estimates of functioning for any one adolescent and lack feasibility because of ethical concerns (e.g., obtaining parent consent to collect peer reports; Card & Hodges, 2008).

Recent work in evidence-based assessment of adolescent social anxiety supports use of a third informant: unfamiliar peer confederates. These individuals "stand in" as unfamiliar peers and interact with adolescents in brief, structured, social interactions designed to elicit anxiety (Anderson & Hope, 2009; Beidel, Rao, Scharfstein, Wong, & Alfano, 2010; Deros et al., 2018). In fact, unfamiliar peer confederate reports yield psychometrically sound data that (a) hold incremental value in predicting adolescent social anxiety concerns relative to parent reports, (b) distinguish adolescents on referral status, and (c) predict adolescents' anxiety-related experiences in social interactions as indexed by both adolescent self-reports of state arousal and observer ratings (Deros et al., 2018; Glenn et al., 2019). Taken together, unfamiliar peer confederates provide rich information about adolescents' nonhome peer contexts and thus represent an ideal third informant to report on adolescent social anxiety. In the present study, we leverage a curated, evidence-based set of informants that consists of parents, adolescents, and unfamiliar peer confederates to assess adolescent social anxiety using the trait-score approach.

## Method

### Participants

We recruited study participants from the Washington, D.C., Maryland, and northern Virginia areas using advertisements posted online (e.g., Craigslist) and flyers posted in local businesses (e.g., cafes, libraries, doctors' offices). Participants responded to one of two posted advertisements: (a) study providing a no-cost social anxiety evaluation for adolescents (i.e., *clinic-referred adolescents*) and (b) nonclinical study on family relationships (i.e., *community control adolescents*). All study procedures described below were held consistent across participants regardless of referral status.

Eligible dyads were required (a) to be fluent in English, (b) to understand the consenting and interview process, and (c) to have an adolescent between 14 and

15 years old whom the parent did not report as having a history of learning or developmental disabilities. The total sample included 127 caregiver-adolescent dyads (43 clinic referred, 84 community control). Adolescents were 14 or 15 years old (mean age = 14.46 years,  $SD = 0.50$ ), and 85 adolescents were girls (66.9%). According to parent report, adolescents' race/ethnicity included (values exceed 100% because multiple categories could be selected) African American/Black (53.5%), European American/White (33.1%), other (7.9%), Hispanic/Spanish/Latino/a (10.2%), Asian American/Asian (5.5%), American Indian (0.8%), or "other" (7.9%). Caregivers included the adolescent's biological mother/father (95.3%) or other caregiver (4.7%). Caregivers reported household income using a scale with 10 categories in \$100 increments. According to this scale, 26.0% of families had a weekly household income of \$500 or less, 22.0% had a weekly household income between \$501 and \$900, and 52.0% had a weekly household income of \$901 or more per week.

In all study analyses, we pooled the clinic-referred and community control samples. This approach mimics displays of dimensionally varying social anxiety in the general population and is consistent with dimensional models of psychopathology (Casey, Oliveri, & Insel, 2014). Furthermore, we selected this approach given prior work indicating that dimensional approaches to assessing psychopathology exhibit greater reliability and validity relative to categorical approaches (Markon, Chmielewski, & Miller, 2011). Nonetheless, we conducted analyses to determine whether clinic-referred and community control groups differed on key demographic characteristics (i.e., adolescent age/gender, adolescent racial/ethnic background, family income, parent's relationship to adolescent, parent's marital status). Given the exploratory nature of these tests, we applied a Bonferroni correction (i.e., 11 tests and thus a corrected  $p$  value of .0045) and observed nonsignificant differences between the clinic-referred and community control samples, which justified our pooled-sample approach.

### Procedure

We received institutional review board approval for all study procedures. Before beginning the study, caregivers provided consent for themselves and their adolescent. Adolescents and caregivers completed a battery of counterbalanced survey measures using Qualtrics software. After completing survey measures, adolescents participated in counterbalanced mock-social-interaction tasks with trained peer confederates. We trained undergraduate research assistants to interact as unfamiliar peer confederates with adolescent participants across mock-social-interaction tasks. After

completion of the study, families received \$100 in monetary compensation for participation.

## Measures

**Survey measures.** Parents completed a demographic questionnaire. Parents, adolescents, and unfamiliar peer confederates completed a battery of survey measures to assess the adolescent's social anxiety. These informants completed measures with parallel survey items; only minor modifications were made to fit each informant's perspective (i.e., "My child" for caregiver measures, "I" for adolescent measures, "The participant" for unfamiliar-peer-confederate measures). As described below, unfamiliar peer confederates completed survey measures after their interactions with adolescents in social interaction tasks.

Our aims involved testing the validity of the trait-score approach, including the incremental validity of this approach above and beyond alternative approaches (i.e., individual reports and composite-score approach). This necessitated use of three different measures of adolescent social anxiety to avoid method confounds in measures used across the various approaches. First, parents, adolescents, and unfamiliar peer confederates completed the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998), a 20-item measure for assessing social anxiety concerns that adolescents may experience during social interactions (example item: "I have difficulty making eye contact with others"). Informants rated items on a scale ranging from 0 (*not at all characteristic or true of me*) to 4 (*extremely characteristic or true of me*). Informants' reports on the SIAS display convergent validity and distinguish adolescents on referral status (Deros et al., 2018; Glenn et al., 2019).

Parents and adolescents also completed the widely used Social Phobia and Anxiety Inventory for Children (SPAIC; Beidel, Turner, & Morris, 1995), a 26-item measure in which informants endorse how often the adolescent feels nervous or scared when in various social scenarios (e.g., public speaking). Several items include "subitems" that require informants to rate the adolescent's social anxiety with different interaction partners (e.g., "adults" vs. "boys or girls his/her age that he/she knows" vs. "boys or girls his/her age that he/she doesn't know"). These subitems were averaged at the item level to create a composite score. Parents and adolescents rated items on a scale ranging from 0 (*never*) to 2 (*always*). Parent and adolescent reports on the SPAIC exhibit convergent and construct validity, distinguish adolescents on referral status, relate to observers' ratings of social skills and anxiety, and are sensitive to treatment response (Beidel et al., 1995; Beidel, Turner, Hamlin, & Morris et al., 2000a; Beidel, Turner, & Morris, 2000).

Parents, adolescents, and unfamiliar peer confederates completed parallel forms of the Social Phobia Scale (SPS; Mattick & Clarke, 1998). The SPS is a 20-item measure for assessing adolescents' fears of being scrutinized by others during routine activities (example item: "I become nervous if I have to write in front of other people"). Informants rated items on a scale ranging from 0 (*not at all characteristic or true of me*) to 4 (*extremely characteristic or true of me*). The SPS demonstrates high internal consistency, discriminates individuals on diagnostic status, and is sensitive to treatment response (Deros et al., 2018; Mattick & Clarke, 1998).

In the present study, we used multi-informant SIAS scores in PCA to identify a social anxiety trait score and parent and adolescent SPAIC reports and unfamiliar peer confederate SPS reports as independent assessments of social anxiety in incremental validity analyses. Furthermore, we used parent, adolescent, and unfamiliar peer confederate SPS reports to create a composite variable in tests of the incremental value of the trait score over and above composite scoring methods.

## Behavioral tasks

**Task descriptions.** Adolescents completed a series of counterbalanced social-interaction tasks with unfamiliar peer confederates adapted from prior research on social anxiety among children, adolescents, and adults (Anderson & Hope, 2009; Beidel, Turner, Hamlin, & Morris, 2000; Beidel et al., 2010). Participants took approximately 20 min to complete these tasks, which included a simulated-social-interaction test (SSIT; adapted from Curran, 1982; Beidel, Turner, Hamlin, & Morris, 2000), unstructured-conversation task (UCT; adapted from Beidel et al., 2010), and impromptu-speech task (IST; adapted from Beidel et al., 2010). Detailed descriptions of these tasks and training of unfamiliar peer confederates are available elsewhere (Deros et al., 2018; Glenn et al., 2019). Across tasks, adolescents interacted with research assistants trained to pose as adolescents. Unfamiliar peer confederates were masked to adolescents' clinical information, including referral status, and had no prior interaction with adolescents. Adolescents' reactions when interacting with these unfamiliar peer confederates predict their reactions when interacting with same-age peers in independent tasks (Karp et al., 2018).

In the SSIT, adolescents participated in a series of five role-playing scenes (e.g., offering/accepting assistance, giving/receiving a compliment) with a gender-matched peer confederate. In each scene, adolescents had two opportunities to speak, and we trained peer confederates to provide two scripted responses. In the UCT, adolescents participated in an unstructured, adolescent-driven, 3-min role-play with a gender-matched

peer confederate. We provided the following instructions to adolescents: "Pretend that you are at a new school and don't know anyone." We trained peer confederates to respond neutrally to adolescents and allow adolescents to lead the conversation. In the IST, adolescents participated in a speech task in which they delivered a speech to a small audience of unfamiliar peers about various topics (e.g., public health, politics). Adolescents were provided a 3-min period of preparation, after which they were provided a 10-min period to complete their speech.

**Adolescent social anxiety behavioral ratings.** We used a behavioral coding scheme developed by Beidel and colleagues (Beidel et al., 2010; Beidel, Turner, Hamlin, & Morris, 2000; Scharfstein, Beidel, Sims, & Finnell, 2011) to obtain independent observers' ratings of adolescent social anxiety in social interaction tasks. Two independent observers used video recordings of social interaction tasks to make macrolevel ratings of adolescent social anxiety on a 5-point scale rating from 1 (*animated*) to 5 (*severe anxiety*). Independent observers made a total of seven ratings (i.e., five SSIT ratings, one UCT rating, one IST rating), and we computed a social anxiety composite rating across tasks. Interrater reliability for the two independent observers' ratings of social anxiety were in the excellent range (average intraclass correlation = .76). For the current study, complete behavioral rating data were available for 105 participants.<sup>1</sup>

### **Data-analytic plan**

**Trait-score identification.** We conducted PCA using procedures described by Kraemer et al. (2003). Specifically, we conducted an unrotated PCA using three "items" (i.e., the total scores of parent, adolescent, and unfamiliar peer confederate SIAS reports). In this respect, our subject-to-item ratio (i.e.,  $127/3 = 42.3:1$ ) falls well above the typical subject-to-item ratios deemed "large" within PCA modeling contexts (e.g., 20:1; see Osborne & Costello, 2004). Within this unrotated PCA, we set the number of components to be extracted to three. As described previously, we examined principal component weights for each informant's report to determine whether we identified trait, context, and perspective scores.

**Prediction of independent criterion variables.** To conduct rigorous tests of the incremental validity of an assessment tool and avoid criterion contamination, independent criterion variables should be used that are completely independent from individual informants' reports (Garb, 2003). Thus, in our tests of the incremental validity of the trait-score approach, we conducted a series of hierarchical linear regressions to evaluate whether the

trait score provides incremental prediction of observed adolescent behavior and referral status. We also conducted secondary analyses to determine whether the context and perspective scores provided incremental prediction of these independent criterion variables.

First, we examined whether the SIAS trait score explained variance in observed adolescent social anxiety over and above the explanatory value of individual informants' reports on independent social anxiety measures. Behavioral observations contribute to evidence-based assessment and clinical decision-making tasks, in particular in assessing behaviors crucial to understanding an individual's presenting problems and measuring change in target behaviors over the course of treatment (Groth-Marnat & Wright, 2016). For example, prior work using the behavioral coding scheme leveraged in the present study has found that observed behavior can be used to estimate treatment response (Beidel, Turner, & Morris, 2000; Beidel et al., 2010; Beidel et al., 2014; Bunnell, Beidel, & Mesa, 2013). Thus, in evaluating the trait-score approach, we selected observed behavior as an independent criterion variable to determine whether the trait score could predict social anxiety observed in key contexts in which adolescents experience social anxiety. We constructed three separate regression models (i.e., one per informant) with observed adolescent social anxiety as the criterion variable. For these models, we entered parent (i.e., SPAIC), adolescent (i.e., SPAIC), and unfamiliar peer confederate (i.e., SPS) reports of adolescent social anxiety in the first step as independent variables and the trait score in the second step as an independent variable. These analyses allowed us to determine whether the trait score provided incremental prediction of observed adolescent social anxiety above and beyond prediction provided by well-established, commonly used, individual informants' reports.

Second, we tested the incremental validity of the SIAS trait score using a series of hierarchical logistic regressions. We examined whether the trait score explained variance in adolescent referral status over and above the explanatory value of individual informants' reports on independent social anxiety measures. For clinical decision-making purposes, it is important for measurement tools to distinguish youths with clinically significant concerns from those without such concerns (e.g., Achenbach & Rescorla, 2001). In fact, researchers commonly use referral status to test the psychometric properties of clinical instruments in research on evidence-based assessment of youth anxiety (De Los Reyes & Makol, 2019; Silverman & Ollendick, 2005). Thus, in evaluating the trait-score approach, we selected referral status as an independent criterion variable to determine whether the trait-score approach

could be used to distinguish youths with and without clinically significant social anxiety concerns. We constructed three separate regression models (i.e., one per informant) with adolescent referral status (0 = community control, 1 = clinic-referred) as the criterion variable. For these models, we entered parent (i.e., SPAIC), adolescent (i.e., SPAIC), and unfamiliar peer confederate (i.e., SPS) reports of adolescent social anxiety in the first step as independent variables and the trait score in the second step as an independent variable. These analyses allowed us to determine whether the trait score provided incremental prediction of adolescent referral status above and beyond prediction provided by well-established and commonly used individual informants' reports.

### **Comparing the trait-score and composite-score approaches.**

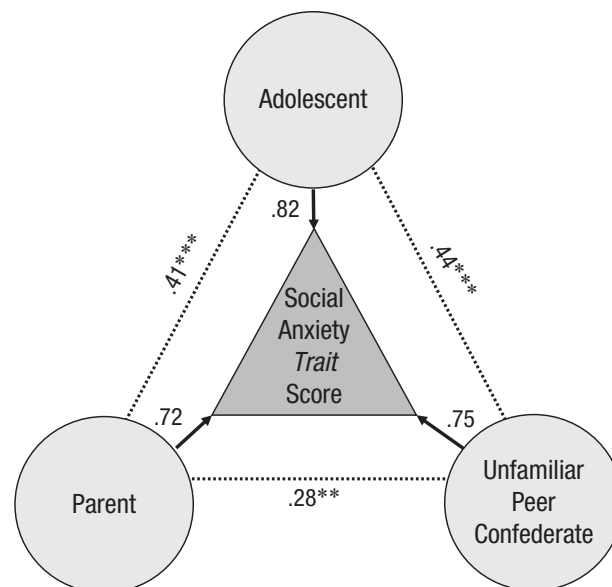
We tested whether the PCA-derived trait score provided incremental prediction of our independent criterion variables over and above a mean-derived composite score. We first computed a social anxiety composite score by creating a variable capturing the average of parent, adolescent, and unfamiliar peer confederate SPS reports ( $M = 21.64$ ,  $SD = 11.28$ ). We selected the SPS to allow for a comparison of the trait-score approach to the composite-score approach using independent measures of the same construct, each with parallel forms (trait score: SIAS; composite score: SPS). We then constructed two separate regression models (i.e., one per criterion variable) with either observed adolescent social anxiety or adolescent referral status (0 = community control, 1 = clinic-referred) as criterion variables. For these models, we entered the SPS composite variable in the first step as an independent variable and the SIAS trait score in the second step as an independent variable. These analyses allowed us to determine whether the PCA-derived trait score provided superior prediction compared with a commonly used integrative method.

For all of our aims, we interpreted statistical significance using a  $p$ -value threshold of  $< .05$ . We inferred magnitudes of effect sizes on the basis of Cohen's (1988) effect-size conventions for the effect size  $r$  (low = .10, moderate = .30, large = .50).

## **Results**

### **Preliminary analyses**

All survey measures displayed acceptable levels of skewness and kurtosis (i.e., scores  $< \pm 2.0$ ) and excellent internal consistency (i.e., all  $\alpha$ s  $\geq .93$ ). We computed bivariate correlations to examine associations within and among multi-informant social anxiety measures used in our study and included this information in Table



**Fig. 2.** Values on dotted lines denote bivariate correlations among parent, adolescent, and unfamiliar peer confederate reports of adolescent social anxiety. Values on solid arrows denote component weights from principal components analysis of parent, adolescent, and unfamiliar peer confederate reports. Asterisks indicate significant correlations between informants' reports (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).

S1 in the Supplemental Material available online. Consistent with prior work (Achenbach et al., 1987; De Los Reyes et al., 2015), correlations for between-informant ratings of adolescent social anxiety were in the small to moderate range (i.e.,  $r$ s = .08–.49). Overall, we observed small to moderate correlations among informants' SIAS reports (Fig. 2), which suggests that we achieved effective triangulation by selecting informants who varied in their context and perspective.

### **Trait-score identification**

In Table 1, we report results from PCA using parent, adolescent, and unfamiliar peer confederate SIAS reports. Consistent with Kraemer et al. (2003), we identified three components that displayed patterns consistent with trait, context, and perspective scores. Specifically, the first component was consistent with a trait score given that loadings for all informants were positive and large in magnitude. The second component was consistent with a context score given that parent and unfamiliar peer confederate reports exhibited contrasting loadings (i.e., reflecting the home vs. nonhome context of informants) and adolescent reports loaded between parent and unfamiliar peer confederate reports (i.e., reflecting both contexts). The third component was consistent with a perspective score given

**Table 1.** PCA of Parent, Adolescent, and Unfamiliar Peer Confederate Reports on the Social Interaction Anxiety Scale (SIAS;  $n = 125$ )

| Informant                   | Component          |                      |                          | Sources of variability in informant's report   |
|-----------------------------|--------------------|----------------------|--------------------------|------------------------------------------------|
|                             | Trait <sup>a</sup> | Context <sup>b</sup> | Perspective <sup>c</sup> |                                                |
| Parent                      | .72                | .64                  | .26                      | Home (context), other (perspective)            |
| Adolescent                  | .82                | -.06                 | -.56                     | Home and nonhome (context), self (perspective) |
| Unfamiliar peer confederate | .75                | -.55                 | .37                      | Nonhome (context), other (perspective)         |

Note: Principal-Component Analysis (PCA) conducted with participants for whom we had full data for parent, adolescent, and unfamiliar peer confederate SIAS reports. Kraemer et al. (2003) described the following criteria for each component: trait (all three informants' reports load strongly and in the same direction), context (informants from different contexts load in opposite directions), and perspective (self-reports load in the opposite direction of observer informants' reports).

<sup>a</sup>Eigenvalue = 1.76; variance attributable to component = 58.5%. <sup>b</sup>Eigenvalue = 0.72; variance attributable to component = 24.1%.

<sup>c</sup>Eigenvalue = 0.52; variance attributable to component = 17.4%.

that parent and unfamiliar peer confederate reports loaded positively and contrasted with adolescent reports (i.e., reflecting the other vs. self-perspective of informants).

### **Prediction of observed adolescent behavior**

In Table 2, we report analyses examining the incremental validity of the SIAS trait score in predicting observed adolescent social anxiety. In the first step of each regression, individual informants' reports of social anxiety predicted observed adolescent social anxiety at moderate to large magnitudes (i.e.,  $\beta$ s = 0.30–0.52). In the second step of each regression, the trait score predicted incremental variance in observed adolescent social anxiety over and above the variance accounted for by individual informants' reports (i.e.,  $\beta$ s = 0.47–0.67). Overall, the trait score accounted for an additional 13% to 29% of the variance in observed social anxiety, which represents moderate to large effects. In additional regression analyses, the context and perspective scores inconsistently predicted incremental variance in observed adolescent social anxiety over and above the variance accounted for by individual informants' reports (i.e., effects ranging from nonsignificant to large in magnitude; see Tables S2 and S3 in the Supplemental Material).

### **Prediction of referral status**

In Table 2, we report analyses examining the incremental validity of the trait score in predicting adolescent referral status. In the first step of each regression, individual informants' reports on independent measures of social anxiety predicted adolescent referral status (i.e., odds ratios [ORs] = 1.03–1.11). In the second step of each regression, the trait score predicted incremental variance in adolescent referral status over and above

the variance accounted for by the individual informants' reports. For every 1-unit increase in the trait score in the second step of regression models, adolescents were approximately 3 to 7 times more likely to be in the clinic-referred group relative to the community control group. In additional regression analyses, the context and perspective scores inconsistently predicted incremental variance in adolescent referral status over and above the variance accounted for by individual informants' reports (i.e., effects ranging from nonsignificant to small in magnitude; see Tables S2 and S3 in the Supplemental Material).

### **Comparing the trait-score and composite-score approaches**

We conducted analyses to determine whether the trait score provided incremental prediction of criterion variables over and above a composite score calculated by averaging the three informants' SPS reports. We found a large and positive correlation between the composite and trait scores ( $r = .88$ ,  $p < .001$ ) and nonsignificant correlations between the composite and context ( $r = -.06$ ,  $p = .48$ ) and composite and perspective ( $r = .004$ ,  $p = .96$ ) scores. This pattern of correlations suggests that although the trait and composite scores show a high degree of overlap, they capture unique information derived from informants' reports.

Using a hierarchical linear regression model, we examined whether the SIAS trait score provided incremental prediction of observed adolescent social anxiety over and above the SPS composite score. In the first step of the regression model, the composite score explained significant variance in observed adolescent social anxiety,  $\beta = 0.54$ ,  $\Delta R^2 = .30$ ,  $\Delta F(1, 101) = 42.55$ ,  $p < .001$ . In the second step of the regression model, the trait score explained incremental variance in observed adolescent social anxiety over and above the

**Table 2.** Hierarchical Regressions Examining the Criterion-Related Validity of SIAS Trait Scores in Predicting Observed Adolescent Social Anxiety and Referral Status

| Step and variable                  | Observed social anxiety ( <i>n</i> = 105) |                |              |         | Adolescent referral status ( <i>N</i> = 127) |           |
|------------------------------------|-------------------------------------------|----------------|--------------|---------|----------------------------------------------|-----------|
|                                    | $\Delta R^2$                              | Total <i>R</i> | <i>b</i>     | $\beta$ | <i>b</i>                                     | <i>OR</i> |
| Parent report                      |                                           |                |              |         |                                              |           |
| Step 1                             | .09**                                     | .30            |              |         |                                              |           |
| SPAIC                              |                                           |                | 0.02 (0.01)  | 0.30**  | 0.11 (0.02)                                  | 1.11***   |
| Step 2                             | .29***                                    | .62            |              |         |                                              |           |
| SPAIC                              |                                           |                | −0.01 (0.01) | −0.10   | 0.07 (0.03)                                  | 1.07**    |
| SIAS, trait score                  |                                           |                | 0.55 (0.08)  | 0.67*** | 0.98 (0.29)                                  | 2.66**    |
| Self-report                        |                                           |                |              |         |                                              |           |
| Step 1                             | .19***                                    | .44            |              |         |                                              |           |
| SPAIC                              |                                           |                | 0.03 (0.01)  | 0.44*** | 0.08 (0.02)                                  | 1.08***   |
| Step 2                             | .19***                                    | .61            |              |         |                                              |           |
| SPAIC                              |                                           |                | −0.01(0.01)  | −0.06   | −0.01 (0.03)                                 | 0.99      |
| SIAS, trait score                  |                                           |                | 0.54 (0.10)  | 0.66*** | 1.35 (0.36)                                  | 3.85***   |
| Unfamiliar peer confederate report |                                           |                |              |         |                                              |           |
| Step 1                             | .27***                                    | .52            |              |         |                                              |           |
| SPS                                |                                           |                | 0.03 (0.004) | 0.52*** | 0.02 (0.01)                                  | 1.03*     |
| Step 2                             | .13***                                    | .63            |              |         |                                              |           |
| SPS                                |                                           |                | 0.01 (0.01)  | 0.22*   | −0.04 (0.02)                                 | 0.96*     |
| SIAS, trait score                  |                                           |                | 0.39 (0.09)  | 0.47*** | 1.88 (0.40)                                  | 6.53***   |

Note: Values in parentheses are standard errors. SPAIC = Social Phobia and Anxiety Inventory for Children (Beidel, Turner, & Morris, 1995); SIAS = Social Interaction Anxiety Scale (Mattick & Clarke, 1998); SPS = Social Phobia Scale (Mattick & Clarke, 1998).

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

composite score,  $\beta = 0.60$ ,  $\Delta R^2 = .08$ ,  $\Delta F(1, 100) = 12.75$ ,  $p < .01$ .

Using a hierarchical logistic regression model, we examined whether the SIAS trait score provided incremental prediction of adolescent referral status over and above the SPS composite score. In the first step of the regression model, the composite score explained significant variance in adolescent referral status,  $b = 0.09$ ,  $OR = 1.10$ ,  $p < .001$ . In the second step of the regression model, the trait score explained significant variance in adolescent referral status over and above the composite score,  $b = 1.45$ ,  $OR = 4.26$ ,  $p < .01$ .

To fully evaluate the incremental validity of the trait score, we conducted additional regression analyses in which we examined whether the SPS composite score provided incremental prediction of observed adolescent social anxiety and referral status over and above the SIAS trait score. In the first step of the regression model, the trait score explained significant variance in observed adolescent social anxiety,  $\beta = 0.61$ ,  $\Delta R^2 = .38$ ,  $\Delta F(1, 101) = 60.84$ ,  $p < .001$ . In the second step of the regression model, the composite score did not explain incremental variance in observed adolescent social anxiety over and above the trait score,  $\beta = 0.01$ ,  $\Delta R^2 = .00$ ,  $\Delta F(1, 100) = .01$ ,  $p = .94$ . In the first step of the regression model,

the trait score explained significant variance in adolescent referral status,  $b = 1.30$ ,  $OR = 3.66$ ,  $p < .001$ . In the second step of the regression model, the composite score did not explain significant variance in adolescent referral status over and above the trait score,  $b = -0.02$ ,  $OR = .99$ ,  $p = .72$ . Overall, these findings support that the trait score, which is derived using PCA, provides incremental prediction of criterion variables over and above well-established individual informants' reports and widely used integrative methods. Specifically, the trait score, which partials out variance from context and perspective (i.e.,  $r$  of 0 between trait score and context and perspective scores) provides a more precise index of the domain being assessed relative to both the composite score and individual informants' reports.

## Discussion

### Main findings

In this study, we rigorously examined the validity of the trait-score approach when applied to adolescent social anxiety assessments. To our knowledge, our study provided the most rigorous and conservative test of the trait-score approach to date through use of

theoretical frameworks and basic science to guide selection of informants, parallel measures across informants, independent criterion variables, and comparisons to an alternative integrative approach. We had three main findings. First, the trait score incrementally predicted observed social anxiety over and above individual informants' reports with moderate- to large-magnitude effects. Beyond individual informants' reports, the trait score explained approximately 10% to 30% more variance in observed adolescent social anxiety. Second, the trait score incrementally predicted adolescent referral status over and above individual informants' reports. Specifically, for every 1-unit increase in the trait score, adolescents were 3 to 7 times more likely to be in the clinic-referred group relative to the community control group. Consistent with Kraemer and colleagues' (2003) focus on the trait score, context and perspective scores failed to predict observed behavior and referral status as robustly and at as high magnitudes as the trait score. Third, we found large magnitude effects for the ability of the trait score to predict observed social anxiety and adolescent referral status over and above a mean-derived composite of informants' reports of social anxiety. In contrast, the composite score did not provide incremental prediction of observed social anxiety and adolescent referral status over and above the trait score. Overall, these findings provide an initial, rigorous demonstration of the trait score's validity, which supports the potential of this approach to improve clinical decision-making in youths' mental health assessments.

### ***Research and theoretical implications***

Kraemer and colleagues (2003) proposed that the trait score represents youth behaviors that vary across contexts and perspectives given that PCA removes unique variance attributable to these components. Researchers commonly apply PCA to individual items in a measure to identify subscales or items that appear to measure common constructs (Tabachnick & Fidell, 2001). Likewise, Kraemer and colleagues applied PCA to individual informants' reports to identify factors that explain variability among these reports. The trait score significantly outperformed the composite score approach in our multi-informant assessment of adolescent social anxiety (the reverse was not true), which supports the idea that the PCA-derived trait score's "triangulation" may provide optimal prediction. This finding suggests that not all integrative strategies are created equal. Although the trait and composite scores were correlated (because we included the same three informants' reports of social anxiety, although on unique measures), the composite score is limited in that it is a simple average of

informants' reports. Using the GPS metaphor, the trait score accounts for the positioning and relative distance of satellites (i.e., informants), whereas the composite score aggregates satellite information without any attention to their source or relative distance. More fundamentally, theory underlying the composite score is in direct contrast not only to theory underlying the trait score but also to a considerable body of research on multi-informant assessments.

Specifically, the composite-score approach assumes that each informant provides an error-prone estimate that when aggregated, points to a "true mean" representation of the domain assessed (Borsboom, 2005; Edgeworth, 1888). This assumption logically extends to collecting reports that achieve maximal *collinearity* with one another, informants for whom reports correspond to a significant extent. In contrast, the trait-score approach assumes that informants' disagreements do not merely reflect measurement error. In fact, they vary systematically and predictably, requiring maximal *orthogonality* among reports (Kraemer et al., 2003). As reviewed previously, research over the past decade supports treating informants' disagreements as reflecting systematic and clinically valuable information (De Los Reyes et al., 2013; De Los Reyes et al., 2015). Overall, our pattern of findings provides strong support for conceptualizing and integrating informants' reports as orthogonal information sources.

Over and above individual informants' reports, the trait score provided incremental prediction of observed adolescent behavior and referral status. These criterion variables differed in measurement method, which provides support for use of the trait score in prediction of both dimensional and discrete clinical indices. As previously mentioned, informants' reports of youths' mental health and particularly youths' internalizing problems are associated with observed behavior at just small to moderate levels. Benchmarked against prior work, the trait score provided more powerful prediction; it accounted for up to 30% more variance in observed behavior beyond variance accounted for by individual informants' reports. Note that we obtained observations of adolescent social anxiety in the context of mock social interactions with unfamiliar peers. The trait score predicted observed anxiety in the nonhome peer context and more strongly than an informant observing adolescent behavior in that same context (i.e., unfamiliar peer confederates). Thus, the trait score outperformed a context-specific informant when predicting context-specific behavior, which suggests that it is a useful tool for understanding youths' mental health concerns globally (i.e., referral status) as well as in specific contexts (i.e., interactions with unfamiliar peers).

## ***Clinical implications***

In practice, low informant convergence is often interpreted as reflecting informant bias or measurement error (De Los Reyes, 2011), which results in suboptimal use of these reports. Specifically, when confronted with informant discrepancies, clinicians often make decisions that “toss out” several informants’ reports in favor of a single informant’s report (Brown-Jacobsen et al., 2011; Hawley & Weisz, 2003; Marsh et al., 2020). Given this leaning to remove reports and the findings reported in this article, we suspect that one’s hypothesis about the meaning underlying informant discrepancies (e.g., they reflect measurement error, one informant provides “better” reports) might result in idiosyncratic decision-making for important clinical tasks such as diagnosis and treatment selection. A large body of research demonstrates that these idiosyncrasies not only commonly occur in clinical work but also that statistical- or actuarial-based decision-making results in relative improvements in accuracy and effectiveness (Grove, Zald, Lebow, Snitz, & Nelson, 2000; Lilienfeld & Lynn, 2014; Meehl, 1954; Youngstrom, Halverson, Youngstrom, Lindheim, & Findling, 2018). Future research is needed to understand how the trait-score approach can move clinical decision-making when using multi-informant reports from an idiosyncratic, dismissive stance on specific information sources to an evidence-based, inclusive stance on multiple information sources.

The trait-score approach can be feasibly implemented in clinical settings and is in line with current evidence-based assessment practices. Estimation of component scores using PCA involves applying a linear transformation to informants’ reports to convert these reports into orthogonal variables. Thus, much like calculation of  $t$  scores in scoring programs for commonly used evidence-based assessments (e.g., Achenbach System of Empirically Based Assessment [ASEBA], Achenbach, 1993; Behavioral Assessment System for Children, Reynolds & Kamphaus, 2004), calculation of the trait score could be programmed into scoring software and included in output designed to enhance clinician interpretation. For example, to enhance the clinical utility of the ASEBA measures, Achenbach’s (1993) manual included core profiles based on cluster analysis and exposition that detailed a potential empirical taxonomy based on the multivariate analyses. McDermott and Weiss’s (1995) validation study of the Adjustment Scales for Children and Adolescents (ASCA) included detailed synopses derived from cluster analysis of participants’  $t$ -score profiles across the ASCA’s six syndromes. Furthermore, the technical manual included instructions for determining which cluster best matched a new case’s profile; these, too, were programmed into spreadsheets

for personal use. More recent efforts indicate that one does not need the infrastructure of a well-established clinical instrument to apply PCA findings to interpreting clinical data. Specifically, to increase dissemination of evidence-based assessment practices in clinical settings, researchers have developed straightforward Excel spreadsheets for entering scores for individual clients to obtain multivariate profiles (Child F.I.R.S.T., 2020; Youngstrom, 2014). Thus, strategies such as these provide proof of concept for the application of the trait-score approach in clinical settings. These strategies can inform the development of paradigms for using the trait score as an efficient integrative tool for multi-informant reports administered at the individual level and thus be commensurate for use in service-delivery settings.

## ***Limitations and future directions***

Our study represented an early-stage illustration of the validity of the trait-score approach. It will be important for future research to determine whether our findings replicate in larger samples that vary in key clinical and demographic characteristics. In our study, we leveraged the trait-score approach at the sample level, similar to use of factor-analytic techniques to arrive at subscales for use in refining measurement tools. Although the development of strategies for applying the trait-score approach to individual cases awaits further study, we expect the current study to inform strategies for use with individual-level clinical decision-making. In particular, we encourage future research that leverages both larger and nationally representative samples and norming procedures that allow trait scores to be interpretable at the individual level.

Kraemer and colleagues (2003) proposed that the trait score, and not the context and perspective scores, provides maximal prediction of clinical variables. Consistent with this proposal, our findings suggest that the context and perspective scores are unlikely to provide robust prediction of independent criterion variables. PCA identifies orthogonal principal components, with the first principal component accounting for the largest amount of item variance. Thus, it is unsurprising that the context and perspective scores displayed inferior prediction of independent criterion variables given that they accounted for less variance in informants’ reports. Note that our findings may be unique given that our study design was optimized for high levels of internal validity. That is, by design, we systematically varied informants’ contexts: Parents and unfamiliar peer confederates had no access to the others’ observations, and thus, we rendered impossible opportunities for cross talk or the exchanging of clinically relevant information between these informants. Despite the likely occurrence of cross talk

in clinical settings (e.g., parent–teacher conferences), a large evidence base supports that informants in these settings often provide discrepant reports and that patterns among these reports provide meaningful clinical information (Achenbach et al., 1987; De Los Reyes et al., 2015). In the trait-score approach, informant cross talk may be conceptualized as unsystematic measurement error, with systematic factors such as informants' contexts and perspectives accounting for more variance among reports. Overall, prior work supports the idea that informant "satellites" can receive some overlapping information about youths' behavior without undermining effective triangulation.

We used unfamiliar peer confederates in the trait-score approach given their access to adolescents' social anxiety in peer contexts, in which parents often have rare opportunity to observe adolescent behavior (Deros et al., 2018). However, these reports are rarely collected in research and clinical settings and may not be feasible to implement in some settings. Future research should rigorously examine the trait-score approach among commonly used informants. Supporting that the trait-score approach will generalize to commonly used informants, meta-analytic work demonstrates that small to moderate convergence is common among parent–youth, parent–teacher, and teacher–youth reports of youths' mental health (Achenbach et al., 1987; De Los Reyes et al., 2015). These levels of convergence approximate those needed to triangulate effectively on a trait score, which suggests that these informants systematically vary in how they rate youths' mental health.

Although the present study focused on social anxiety, we expect findings to generalize to rigorous research applying the trait-score approach to other problem domains. Future research should thoughtfully apply the mix-and-match criterion to identify optimal informants and multimodal indices in mental health assessments. For example, the trait-score approach may improve the information derived in externalizing problem assessments. Given the importance of understanding externalizing problems across key contexts (i.e., home, school), most research on the interpretability of multi-informant reports and cross-contextual variations in mental health concerns focuses on externalizing problems (e.g., disruptive behavior, attentional concerns; Drabick, Bubier, Chen, Price, & Lanza, 2011; Strickland, Hopkins, & Keenan, 2012). Thus, the trait-score approach is well suited to externalizing problem assessments.

We also expect that the issues discussed in this article, which focused on integrating multiple information sources, will be relevant to basic research on psychopathology. For example, the Research Domain Criteria (RDoC) framework seeks to develop and validate dimensional constructs derived from multiple units of

analysis (e.g., subjective reports, behaviors, physiology) in an effort to improve clinical decision-making (Casey et al., 2014). RDoC approaches rely on evidence-based assessment of key constructs as well as data-analytic models for integrating these multivariate data. Recent work highlights the challenges of accomplishing these tasks (Youngstrom et al., 2018). Mental health concerns can be particularly challenging to assess not just because of low convergence among informants' reports but also because of the low convergence generally among units of analysis within and across functional domains relevant to externalizing disorders (e.g., social processes, regulatory functioning; Beauchaine & Hinshaw, 2020). We anticipate that research in this area will benefit from conceptualizing disparate information sources as satellites collecting unique information that varies for systematic and predictable reasons and, consequently, will interpret low measurement convergence as a tool for optimizing prediction (De Los Reyes, Drabick, Makol, & Jakubovic, 2020). Overall, when integrated with contemporary conceptual models informing research on multi-informant assessments (OTM; De Los Reyes et al., 2013), the theory and method of the trait-score approach advanced by Kraemer and colleagues (2003) has the potential to transform widely used practices in multi-informant assessments as well as to inform clinical developmental research on a broad array of other units of analysis.

## Transparency

*Action Editor:* Erin B. Tone

*Editor:* Scott O. Lilienfeld

### Author Contributions

A. De Los Reyes designed the study. B. A. Makol, N. Qasmieh, and A. De Los Reyes executed the study. N. Qasmieh assisted with data analyses and B. A. Makol completed data analyses. B. A. Makol and A. De Los Reyes wrote the manuscript. E. A. Youngstrom, S. J. Racz, L. E. Glenn, and N. Qasmieh collaborated in editing the manuscript. All the authors approved the final version of the manuscript for submission.

### Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

### Funding

This work was supported by the Institute of Education Sciences.

### Open Practices

This article is based on archival data that is part of a larger study for which data collection is ongoing. Once data collection is complete, the authors intend to make the data set from the larger study available for public use. A full description of the materials and procedures used to collect data for this article is available at <https://osf.io/mcdth>. The complete Open Practices Disclosure for this article can be

found at <http://journals.sagepub.com/doi/suppl/10.1177/2167702620924439>. More information about the Open Practices badges can be found at <https://www.psychologicalscience.org/publications/badges>.

### ORCID iDs

Bridget A. Makol  <https://orcid.org/0000-0002-7627-2614>  
Eric A. Youngstrom  <https://orcid.org/0000-0003-2251-6860>

### Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/2167702620924439>

### Note

1. We conducted  $\chi^2$  analyses to determine whether participants with and without behavioral rating data differed on key demographic characteristics (i.e., adolescent age/gender, family income, parent's relationship to adolescent, parent's marital status). We did not conduct  $\chi^2$  analyses for race/ethnicity for these two groups given that doing so would require comparisons of groups with cell sizes including fewer than 5 participants. Given the exploratory nature of these tests, we applied a Bonferroni correction (i.e., five tests and thus a corrected  $p$  value of .01). We observed nonsignificant differences between participants with and without behavior rating data available.

### References

- Achenbach, T. M. (1993). *Empirically based taxonomy: How to use syndromes and profile types derived from the CBCL/4-18, TRF, and YSR*. Burlington: University of Vermont Department of Psychiatry.
- 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. doi:10.1037/0033-2909.101.2.213
- Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment*. Burlington: University of Vermont, Research Center for Children, Youth & Families.
- Anderson, E. R., & Hope, D. A. (2009). The relationship among social phobia, objective and perceived physiological reactivity, and anxiety sensitivity in an adolescent population. *Journal of Anxiety Disorders*, 23, 18–26. doi:10.1016/j.janxdis.2008.03.011
- Becker, S. P., Luebbe, A. M., Fite, P. J., Stoppelbein, L., & Greening, L. (2014). Sluggish cognitive tempo in psychiatrically hospitalized children: Factor structure and relations to internalizing symptoms, social problems, and observed behavioral dysregulation. *Journal of Abnormal Child Psychology*, 42, 49–62. doi:10.1007/s10802-013-9719-y
- Becker-Haimes, E. M., Jensen-Doss, A., Birmaher, B., Kendall, P. C., & Ginsburg, G. S. (2017). Parent–youth informant disagreement: Implications for youth anxiety treatment. *Clinical Child Psychology and Psychiatry*, 23, 42–56. doi:10.1177/1359104516689586
- Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., . . . Mandell, D. S. (2015). Free, brief, and validated: Standardized instruments for low-resource mental health settings. *Cognitive and Behavioral Practice*, 22, 5–19. doi:10.1016/j.cbpra.2014.02.002
- Beidel, D. C., Alfano, C. A., Kofler, M. J., Rao, P. A., Scharfstein, L., & Sarver, N. W. (2014). The impact of social skills training for social anxiety disorder: A randomized controlled trial. *Journal of Anxiety Disorders*, 28, 908–918. doi:10.1016/j.janxdis.2014.09.016
- Beidel, D. C., Rao, P. A., Scharfstein, L., Wong, N., & Alfano, C. A. (2010). Social skills and social phobia: An investigation of DSM-IV subtypes. *Behaviour Research and Therapy*, 48, 992–1001. doi:10.1016/j.brat.2010.06.005
- Beidel, D. C., Turner, S. M., Hamlin, K., & Morris, T. L. (2000). The Social Phobia and Anxiety Inventory for Children (SPAI-C): External and discriminative validity. *Behavior Therapy*, 31, 75–87. doi:10.1016/S0005-7894(00)80005-2
- Beidel, D. C., Turner, S. M., & Morris, T. L. (1995). A new inventory to assess childhood social anxiety and phobia: The Social Phobia and Anxiety Inventory for Children. *Psychological Assessment*, 7, 73–79. doi:10.1037/1040-3590.7.1.73
- Beidel, D. C., Turner, S. M., & Morris, T. L. (2000). Behavioral treatment of childhood social phobia. *Journal of Consulting and Clinical Psychology*, 68, 1072–1080. doi:10.1037/0022-006X.68.6.1072
- Beauchaine, T. P., & Hinshaw, S. P. (2020). RDoC and psychopathology among youth: Misplaced assumptions and an agenda for future research. *Journal of Clinical Child & Adolescent Psychology*, 49, 322–340. doi:10.1080/15374416.2020.1750022
- Borsboom, D. (2005). *Measuring the mind*. New York, NY: Cambridge University Press.
- Brown-Jacobsen, A. M., Wallace, D. P., & Whiteside, S. P. (2011). Multimethod, multi-informant agreement, and positive predictive value in the identification of child anxiety disorders using the SCAS and ADIS-C. *Assessment*, 18, 382–392. doi:10.1177/1073191110375792
- Bunnell, B. E., Beidel, D. C., & Mesa, F. (2013). A randomized trial of attention training for generalized social phobia: Does attention training change social behavior? *Behavior Therapy*, 44, 662–673. doi:10.1016/j.beth.2013.04.010
- Card, N. A., & Hodges, E. V. E. (2008). Peer victimization among school children: Correlations, causes, consequences, and considerations in assessment and intervention. *School Psychology Quarterly*, 23, 451–461. doi:10.1037/a0012769
- Cartwright-Hatton, S., Hodges, L., & Porter, J. (2003). Social anxiety in childhood: The relationship with self and observer rated social skills. *Journal of Child Psychology and Psychiatry*, 44, 737–742. doi:10.1111/1469-7610.00159
- Casey, B. J., Oliveri, M. E., & Insel, T. (2014). A neurodevelopmental perspective on the research domain criteria (RDoC) framework. *Biological Psychiatry*, 76, 350–353. doi:10.1016/j.biopsych.2014.01.006
- Child F.I.R.S.T. (2020). *Focus on innovation and redesign in systems and treatment*. Retrieved from <https://www.childfirst.ucla.edu/resources/>

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Curran, J. P. (1982). A procedure for the assessment of social skills: The Simulated Social Interaction Test. In J. P. Curran & P. M. Monti (Eds.), *Social skills training: A practical handbook for assessment and treatment* (pp. 348–373). New York, NY: Guilford.
- De Los Reyes, A. (2011). More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 40, 1–9. doi:10.1080/15374416.2011.533405
- De Los Reyes, A., Alfano, C. A., Lau, S., Augenstein, T. M., & Borelli, J. L. (2016). Can we use convergence between caregiver reports of adolescent mental health to index severity of adolescent mental health concerns? *Journal of Child and Family Studies*, 25, 109–123. doi:10.1007/s10826-015-0216-5
- De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141, 858–900. doi:10.1037/a0038498
- De Los Reyes, A., Drabick, D. A. G., Makol, B. A., & Jakubovic, R. J. (2020). Introduction to the special section: The Research Domain Criteria's units of analysis and cross-unit correspondence in youth mental health research. *Journal of Clinical Child and Adolescent Psychology*, 49, 279–296. doi:10.1080/15374416.2020.1738238
- De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*, 37, 637–652. doi:10.1007/s10802-009-9307-3
- De Los Reyes, A., & Langer, D. A. (2018). Assessment and the *Journal of Clinical Child and Adolescent Psychology's* Evidence Base Updates Series: Evaluating the tools for gathering evidence. *Journal of Clinical Child & Adolescent Psychology*, 47, 357–365. doi:10.1080/15374416.2018.1458314
- De Los Reyes, A., & Makol, B. A. (2019). Evidence-based assessment. In T. H. Ollendick, L. Farrell, & P. Muris (Eds.), *Innovations in CBT for childhood anxiety, OCD, and PTSD: Improving access and outcomes* (pp. 28–51). New York, NY: Cambridge.
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123–149. doi:10.1146/annurev-clinpsy-050212-185617
- Deros, D. E., Racz, S. J., Lipton, M. F., Augenstein, T. M., Karp, J. N., Keeley, L. M., . . . De Los Reyes, A. (2018). Multi-informant assessments of adolescent social anxiety: Adding clarity by leveraging reports from unfamiliar peer confederates. *Behavior Therapy*, 49, 84–98. doi:10.1016/j.beth.2017.05.001
- Dirks, M. A., De Los Reyes, A., Briggs-Gowan, M., Cella, D., & Wakschlag, L. S. (2012). Annual Research Review: Embracing not erasing contextual variability in children's behavior-theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 53, 558–574. doi:10.1111/j.1469-7610.2012.02537.x
- Doctoroff, G. L., & Arnold, D. H. (2004). Parent-rated externalizing behavior in preschoolers: The predictive utility of structured interviews, teacher reports, and classroom observations. *Journal of Clinical Child and Adolescent Psychology*, 33, 813–818. doi:10.1207/s15374424jccp3304\_16
- Drabick, D. A. G., Bubier, J., Chen, D., Price, J., & Lanza, I. H. (2011). Source-specific oppositional defiant disorder among inner-city children: Prospective prediction and moderation. *Journal of Clinical Child and Adolescent Psychology*, 40, 23–35. doi:10.1080/15374416.2011.533401
- Edgeworth, F. Y. (1888). The statistics of examinations. *Journal of the Royal Statistical Society*, 51, 598–635.
- Garb, H. N. (2003). Incremental validity and the assessment of psychopathology in adults. *Psychological Assessment*, 15, 508–520. doi:10.1037/1040-3590.15.4.508
- Gazelle, H., Workman, J. O., & Allan, W. (2010). Anxious solitude and clinical disorder in middle childhood: Bridging developmental and clinical approaches to childhood social anxiety. *Journal of Abnormal Child Psychology*, 38, 1–17. doi:10.1007/s10802-009-9343-z
- Glenn, L. E., Keeley, L. M., Szollos, S., Okuno, H., Wang, X., Rausch, E., . . . De Los Reyes, A. (2019). Trained observers' ratings of adolescents' social anxiety and social skills within controlled, cross-contextual social interactions with unfamiliar peer confederates. *Journal of Psychopathology and Behavioral Assessment*, 41, 1–15. doi:10.1007/s10862-018-9676-4
- Groth-Marnat, G., & Wright, A. J. (2016). *Handbook of psychological assessment* (6th ed.). Hoboken, NJ: Wiley.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19–30. doi:10.1037/1040-3590.12.1.19
- Hartley, A. G., Zakriski, A. L., & Wright, J. C. (2011). Probing the depths of informant discrepancies: Contextual influences on divergence and convergence. *Journal of Clinical Child & Adolescent Psychology*, 40, 54–66. doi:10.1080/15374416.2011.533404
- Hawley, K. M., & Weisz, J. R. (2003). Child, parent and therapist (dis)agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *Journal of Consulting and Clinical Psychology*, 71, 62–70. doi:10.1037/0022-006X.71.1.62
- Henry, D. B., & Metropolitan Area Child Study Research Group. (2006). Associations between peer nominations, teacher ratings, self-reports, and observations of malicious and disruptive behavior. *Assessment*, 13, 241–252. doi:10.1177/1073191106287668
- Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology*, 21, 143–150.
- Hofmann, S. G., Albano, A. M., Heimberg, R. G., Tracey, S., Chorpita, B. F., & Barlow, D. H. (1999). Subtypes of social phobia in adolescents. *Depression and Anxiety*, 9, 15–18. doi:10.1002/(SICI)1520-6394(1999)9:1 < 15::AID-DA2 > 3.0.CO;2-6

- Houts, R. M., Caspi, A., Pianta, R. C., Arseneault, L., & Moffitt, T. E. (2010). The challenging pupil in the classroom: The effect of the child on the teacher. *Psychological Science*, 21, 1802–1810. doi:10.1177/0956797610388047
- Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, 3, 29–51. doi:10.1146/annurev.clinpsy.3.022806.091419
- Jones, J. D., Boyd, R. C., Calkins, M. E., Ahmed, A., Moore, T. M., Barzilay, R., . . . Gur, R. E. (2019). Parent-adolescent agreement about adolescents' suicidal thoughts. *Pediatrics*, 143, 1–12. doi:10.1542/peds.2018-1771.
- Karp, J., Makol, B. A., Keeley, L. M., Qasmieh, N., Deros, D. E., Racz, S. J., . . . De Los Reyes, A. (2018). Convergent, incremental, and criterion-related validity of multi-informant assessments of adolescents' fears of negative and positive evaluation. *Clinical Psychology and Psychotherapy*, 25, 217–230. doi:10.1002/cpp.2156
- 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. doi:10.1176/appi.ajp.160.9.1566
- Kwon, K., Kim, E. M., & Sheridan, S. M. (2012). A contextual approach to social skills assessment in the peer group: Who is the best judge? *School Psychology Quarterly*, 27, 121–133. doi:10.1037/a0028696
- Lerner, M. D., De Los Reyes, A., Drabick, D. G., Gerber, A. H., & Gadow, K. D. (2017). Informant discrepancy defines discrete, clinically useful autism spectrum disorder subgroups. *Journal of Child Psychology and Psychiatry*, 58, 829–839. doi:10.1111/jcpp.12730
- Lilienfeld, S. O., & Lynn, S. J. (2014). Errors/biases in clinical decision making. In *The encyclopedia of clinical psychology*. doi:10.1002/9781118625392.wbecp567
- Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants of childhood psychopathology. *Journal of Clinical Child Psychology*, 19, 136–143. doi:10.1207/s15374424jccp1902\_5
- Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology*, 1, 317–337. doi:10.1017/S095457940000050X
- Makol, B. A., De Los Reyes, A., Ostrander, R., & Reynolds, E. K. (2019). Parent-youth divergence (and convergence) in reports of youth internalizing problems in psychiatric inpatient care. *Journal of Abnormal Child Psychology*, 47, 1677–1689. doi:10.1007/s10802-019-00540-7
- Makol, B. A., & Polo, A. J. (2018). Parent-child endorsement discrepancies among youth at chronic-risk for depression. *Journal of Abnormal Child Psychology*, 46, 1977–1088. doi:10.1007/s10802-017-0360-z
- Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*, 137, 856–879. doi:10.1037/a0023678.
- Marsh, J. K., De Los Reyes, A., & Lilienfeld, S. O. (2018). Leveraging the multiple lenses of psychological science to inform clinical decision making: Introduction to the special section. *Clinical Psychological Science*, 6, 167–176. doi:10.1177/2167702617736853
- Marsh, J. K., Zeveney, A., & De Los Reyes, A. (2020). Informant discrepancies in judgments about change during mental health treatments. *Clinical Psychological Science*, 8, 318–332. doi:10.1177/2167702619894905
- Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36, 455–470. doi:10.1016/S0005-7967(97)10031-6
- McDermott, P. A., & Weiss, R. V. (1995). A normative typology of healthy, subclinical, and clinical behavior styles among American children and adolescents. *Psychological Assessment*, 7, 162–170. doi:10.1037/1040-3590.7.2.162
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis, MN: University of Minnesota Press.
- Noordhof, A., Oldehinkel, A. J., Verhulst, F. C., & Ormel, J. (2008). Optimal use of multi-informant data on co-occurrence of internalizing and externalizing problems: The TRAILS study. *International Journal of Methods in Psychiatric Research*, 17, 174–183.
- Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research, and Evaluation*, 9(1), Article 11. doi:10.7275/ktzq-jq66
- Perren, S., Stadelmann, S., Burgin, D., & Von Klitzing, K. A. I. (2006). Associations between behavioral/emotional difficulties in kindergarten children and the quality of their peer relationships. *Journal of the American Academy of Child & Adolescent Psychiatry*, 45, 867–876. doi:10.1097/01.chi.0000220853.71521.cb
- Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children* (2nd ed.). Circle Pines, MN: American Guidance Service.
- Scharfstein, L. A., Beidel, D. C., Sims, V. K., & Fennell, L. R. (2011). Social skills deficits and vocal characteristics of children with social phobia or Asperger's disorder: A comparative study. *Journal of Abnormal Child Psychology*, 39, 865–875. doi:10.1007/s10802-011-9498-2
- Silverman, W. K., & Ollendick, T. H. (2005). Evidence-based assessment of anxiety and its disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 34, 380–411. doi:10.1207/s15374424jccp3403\_2
- Smetana, J. G. (2008). "It's 10 o'clock: Do you know where your children are?" Recent advances in understanding parental monitoring and adolescents' information management. *Child Development Perspectives*, 2, 19–25. doi:10.1111/j.1750-8606.2008.00036.x
- Strickland, J., Hopkins, J., & Keenan, K. (2012). Mother-teacher agreement on preschoolers' symptoms of ODD and CD: Does context matter? *Journal of Abnormal Child Psychology*, 40, 933–943. doi:10.1007/s10802-012-9622-y
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn and Bacon.
- Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B. L., Keenan, K., . . . Carter, A. S. (2008).

- Observational assessment of preschool disruptive behavior, part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child & Adolescent Psychiatry*, 47, 632–641. doi:10.1097/CHI.0b013e31816c5c10
- Winsler, A., & Wallace, G. L. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education and Development*, 13, 41–58. doi:10.1207/s15566935eed1301\_3
- Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, 39, 204–221. doi:10.1093/jpepsy/jst062
- Youngstrom, E. A., Halverson, T. F., Youngstrom, J. K., Lindhiem, O., & Findling, R. L. (2018). Evidence-based assessment from simple clinical judgments to statistical learning: Evaluating a range of options using pediatric bipolar disorder as a diagnostic challenge. *Clinical Psychological Science*, 6, 243–265. doi:10.1177/2167702617741845