Running Head: MULTI-INFORMANT, MULTIMETHOD ASSESSMENT

Interpreting Convergences and Divergences in Multi-Informant, Multi-Method Assessment

Andres De Los Reyes[1] **Bridget A. Makol**[1]

[1]University of Maryland at College Park

Author Notes

Andres De Los Reyes and Bridget A. Makol, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland, College Park, MD.

Efforts by the first and second authors were supported by a grant from the Institute of Education Sciences (R324A180032).

Correspondence regarding this manuscript should be addressed to 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; Office: 301-405-7049; Fax: 301-314-9566; E-mail: adlr@umd.edu; Twitter: @JCCAP_Editor

CITATION

De Los Reyes, A., & Makol, B. A. (2021). Interpreting convergences and divergences in multi-informant, multi-method assessment. In J. Mihura (Ed.), *The Oxford handbook of personality and psychopathology assessment*. (2nd ed.). Oxford. (in press)

Abstract

Individuals often vary in how they display signs and symptoms of personality disturbances and psychopathology. As such, comprehensive assessments of these signs and symptoms ought to capture how they manifest within and across time and contexts (e.g., community, home, work). Contexts may vary in the degree to which they influence personality and psychopathology concerns. Thus, practitioners and researchers often rely on multiple informants and methodologies to characterize intervention targets, monitor intervention progress, and inform the selection of evidence-based services. Across research teams and assessment domains, researchers commonly observe diverging outcomes among assessments, depending on the informant and/or method. We review theory and research that demonstrates how patterns of divergence observed in multi-informant, multi-method assessments represent valid individual differences in clinical presentations. This divergence may inform interpretations of personality and psychopathology functioning clinically, and we advance a research agenda to improve their use when assessing and diagnosing adult personality and psychopathology.

Keywords: informant discrepancies; multiple informants; Operations Triad Model; personality

Interpreting Convergences and Divergences in Multi-Informant, Multi-Method Assessment

Imagine you are a clinician administering an assessment to an adult client for whom the referral question involved evaluating signs and symptoms of avoidant personality disorder and associated pathological personality traits. As part of this evaluation and consistent with the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association [APA], 2013), you administer measures developed to assess for key criteria, including Criterion A domains of personality functioning (i.e., identity, self-direction, empathy, intimacy), and Criterion B domains of pathological traits (i.e., anxiousness, withdrawal, anhedonia, intimacy avoidance). You solicit reports on these measures from the client, and with your client's consent, a report from a significant other in their life, namely their coworker. As the assessing clinician and following your careful observation of the client, you also complete measures of the client's functioning. After gathering this information, you observe disagreements among the reports. The coworker reported concerns with self-direction for which the client's self-reports did not corroborate, and the client self-reported empathy concerns for which their coworker's reports did not corroborate. When completing your own reports of the client, you encountered uncertainty given that your observations revealed concerns in other domains of functioning beyond those endorsed by the client and coworker. Within this evaluation, what might account for the disagreements among reports? Does the client lack sufficient insight into their concerns to provide psychometrically sound reports? Does the coworker know the client well enough to provide psychometrically sound reports? Do these disagreements affect your own impressions of the client, and how do you integrate all of these data to make precise, accurate clinical decisions (e.g., forming diagnoses, planning treatment)?

In this chapter, we review research on divergent assessment outcomes observed in multi-informant, multi-method assessments of mental health (hereafter collectively referred to as "comprehensive mental health assessments"), and strategies for interpreting and integrating these divergent assessments. We focus on several areas of work examining divergence among information sources in comprehensive mental health assessments. Specifically, we provide an overview of research on divergent assessments as they manifest in adult personality and psychopathology assessments, and importantly, across the myriad assessment literatures where researchers have documented their presence. We leverage work within and across these literatures as the foundation for covering three key issues germane to interpreting divergent outcomes in comprehensive mental health assessments. First, we discuss issues relevant to assumptions regarding how adults display signs and symptoms of personality and psychopathology as well as the veracity of specific informants' reports (e.g., clients and collateral informants). Second, we describe conceptual and measurement models designed to facilitate understanding, interpreting, and integrating comprehensive mental health assessment data, and review their evidentiary support. Third, we highlight the research and clinical implications of these models and outline directions for future research.

The Ubiquitous Presence of Divergent Assessment Outcomes

The literature on divergent assessment outcomes in comprehensive assessments of adult mental health presents both challenges and opportunities. We highlight each of these challenges and opportunities in Table 1. Throughout the chapter, we leverage Table 1 to organize our synthesis of the literature and provide recommendations for future research. The overarching theme to which we draw your attention is that the literature on divergent information sources in comprehensive assessments of adult mental health is relatively nascent. Indeed, clinical assessments of adult personality and

psychopathology have historically consisted of a limited number of assessment methodologies and informants, namely self-report surveys or checklists, and methods that rely heavily on self-report (e.g., interviews and clinician ratings; see Achenbach, 2006; Hunsley & Mash, 2007). This reliance on self-reports exists despite research demonstrating that self-reports of adult functioning, and in particular personality functioning, commonly display low-to-moderate levels of correspondence with other modalities including collateral reports, clinician diagnoses, implicit measures, and behavioral observations (Ganellen, 2007; Klonsky et al., 2002; McCrae et al., 2004). Nevertheless, only in the last 15 to 20 years have researchers begun to pay considerable attention to describing levels and moderators of divergent assessment outcomes in clinical work with adults (e.g., Klonsky et al., 2002; Oltmanns & Turkheimer, 2009; Samuel et al., 2016, 2018). We say "only" because, in contrast to clinical work with adults, comprehensive mental health assessments have formed a key part of "best practices" in clinical work with children and adolescents for several decades (e.g., Achenbach, 2017; De Los Reyes, 2011; Hunsley & Mash, 2007). Consequently, one can trace the study of divergent outcomes in child and adolescent assessments back to the 1950s (Lapouse & Monk, 1958), and up through the current state of a literature that consists of over 400 studies (for reviews, see Achenbach et al., 1987; De Los Reyes et al., 2015, 2019a).

Beyond challenges, in this chapter we highlight concrete opportunities for future research on comprehensive assessments of adult mental health. By "opportunities," we mean crucial gaps in the literature on these assessments that, if filled, could revolutionize how practitioners and researchers collect, interpret, and integrate comprehensive assessment data. Along these lines, in Table 1 we provide a "top 25" list of representative publications from the child and adolescent assessment literature, designed to facilitate pursuit of the opportunities that we discuss. We can draw on these lessons learned from other assessment literatures, in large part, because the level

of divergent assessment outcomes in comprehensive mental health assessments is remarkably consistent within and across literatures. In fact, multiple lines of research focused on advancing our understanding of comprehensive assessments of child and adolescent mental health also apply to understanding comprehensive assessments of adults.

For instance, consider that, in the pioneering meta-analysis of correspondence in cross-informant reports of adult mental health conducted by Achenbach and colleagues (2005), the mean cross-informant correlations observed for adult internalizing problems (r = .43) and externalizing problems (r = .44) were consistent with the mean cross-informant estimates for internalizing problems (r = .32) and externalizing problems (r = .41) observed in Achenbach and colleagues' (1987) pioneering meta-analysis of 119 studies of correspondence in child and adolescent assessments. More recent meta-analytic studies in the child and adolescent assessment literature support the robust, low-to-moderate levels of correspondence first observed by Achenbach and colleagues (1987). In fact, the overall level of cross-informant correspondence in reports of child and adolescent mental health reported by Achenbach and colleagues (1987) of .28 was *identical* to the overall estimate (.28) observed in an independent meta-analysis of 341 studies on child and adolescent correspondence published between 1989 and 2014 (De Los Reyes et al., 2015).

Interestingly, the cross-informant correspondence estimates observed in Achenbach and colleagues (1987) and De Los Reyes and colleagues (2015) are within the range of magnitudes of correlations seen in meta-analytic reviews of multi-modal assessments of personality functioning (e.g., comparisons of Rorschach variables with independent estimates of personality functioning; Mihura et al., 2013) and in reviews of assessment literatures more generally (e.g., Hemphill, 2003). Further, a series of additional meta-analyses in the child and adolescent

assessment literature cement the replicability of these low-to-moderate levels of correspondence (Table 1). Indeed, we consistently see these low-to-moderate levels of correspondence manifest across meta-analyses of assessments of such varied domains as autism (Stratis & Lecavalier, 2015), parenting (Hou et al., 2019; Korelitz & Garber, 2016), and social competence (Renk & Phares, 2004). These low-to-moderate levels of correspondence also robustly manifest in meta-analyses of assessments conducted across various settings, such as assessments of maltreated youth (Romano et al., 2018) and assessments conducted across cultures (De Los Reyes et al., 2019a).

As we noted in recent work (De Los Reyes et al., 2020), the replicability of low-to-moderate levels of cross-informant correspondence rivals the replicability of placebo effects (cf. Ashar et al., 2017). The replicability of divergent assessment outcomes allows for the development of assessment paradigms, interpretive strategies, and conceptual and measurement models with broad applicability to assessments of not only multiple mental health domains but also assessments conducted across various settings (e.g., community, educational, foster care, inpatient and outpatient facilities) and developmental periods (e.g., preschool, middle childhood, adolescence, emerging-, middle-, and late-adulthood). Further, work on issues surrounding divergent assessment outcomes as they manifest outside comprehensive assessments of adult mental health may facilitate addressing long-standing challenges, namely with questions surrounding the apparent mismatch between divergent assessment outcomes and the conceptualization, classification, and diagnosis of adult personality and psychopathology.

Assumptions about the Phenomenology of Personality and Psychopathology

Perhaps the most vexing issue confronting the interpretability of divergent outcomes in comprehensive assessments of adult mental health lies in how practitioners and researchers typically

conceptualize domains of personality and psychopathology. In particular, for personality disorders, as well as multiple DSM-5 disorders that fall across the internalizing and externalizing spectra (e.g., social anxiety, attention-deficit/hyperactivity disorder, substance use disorders), signs and symptoms ought to manifest invariantly across contexts and time (APA, 2013). Not surprisingly, many have typically viewed the low levels of cross-informant correspondence in comprehensive assessments of adult mental health as a "problem" requiring a methodological solution (for reviews, see De Los Reyes et al., 2013a, 2019a). That is, if the very constructs one measures should manifest consistently across contexts and time, why is it that reports from informants tasked to provide information about these constructs fail to display relatively high levels of correspondence with one another? When this occurs, a seemingly logical question arises: What is "wrong" with the informants?

Implications for Measurement Assumptions: Biased Informants and "Optimal" Informants

One of the most commonly discussed interpretations of low cross-informant correspondence is that one or more of the informants provides biased or inaccurate reports (e.g., Roberts & Caspi, 2001). In particular, and as stipulated within the DSM-5 (APA, 2013), many have called into question the degree to which clients have sufficient insight to provide adequate self-reports (e.g., Alexander et al., 2017; Balsis et al., 2015; Clark et al., 1997). In fact, assumptions about validity issues with self-reports due to biases, low insight, and purposeful misrepresentation linked to the very problems being assessed appears to be more pronounced for personality symptoms than most other forms of psychopathology (Ganellen, 2007; Huprich et al., 2011). Given that much of the literature on personality disorders relies almost exclusively on self-report (e.g., Oltmanns & Turkheimer, 2009), the relatively low levels of correspondence between self-reports, clinicians, and other informants (e.g., coworkers, friends, romantic partners) has prompted concerns about the veracity of research findings on adult mental health. As evidence of this, concerns about the veracity of multi-informant reports have

occurred alongside a cottage industry of approaches for dealing with low cross-informant correspondence that largely stem from the "methodological confound" interpretation of this phenomenon. These include testing cross-informant differences in the psychometric properties of reports (e.g., Balsis et al., 2015) and identifying the "optimal" or most psychometrically sound informant (e.g., Galione & Oltmanns, 2013; Vazire, 2010; Vazire & Carlson, 2011).

One can identify similar interpretations regarding the veracity of informants' reports in research on adult psychopathology. For example, and much like clients presenting with personality disorders, those presenting with symptoms of psychosis are often thought to lack sufficient insight to provide sound reports of their mental health (e.g., Birchwood et al., 1994). As another example, consider that interpretative biases or evaluative concerns are core features of diagnostic criteria for anxiety and mood disorders (APA, 2013). Clients diagnosed with social anxiety disorder may present with fears of evaluation by others and, in particular, unfamiliar individuals (e.g., Bögels et al., 2010). As such, low cross-informant correspondence may be attributable to clients downplaying their concerns to "look good" in front of assessors (see also De Los Reyes et al., 2015). Similarly, clients diagnosed with mood disorders often present with negatively biased interpretations of themselves, the world, and others (e.g., Richters, 1992) and, as such, they may attend to, encode, recall, and thus provide reports about negative behaviors (e.g., mood symptoms) to a greater degree than positive or neutral behaviors (see also Youngstrom et al., 1999). Many self-reports used to measure mental health among individuals meeting criteria for mental health diagnoses often exhibit sound psychometric properties, as indexed by well-established interpretive guidelines (De Los Reyes & Langer, 2018; Hunsley and Mash, 2008, 2018; Youngstrom et al., 2017). Nevertheless, interpretations regarding rater biases result in a serious question: Can individuals with identified psychopathology provide psychometrically sound reports about their own mental health?

More of the Same: Assumptions in Other Assessment Literatures

Similar to the robust presence of divergent assessment outcomes across literatures, research on comprehensive assessments of adult mental health is in good company with other literatures in terms of casting doubt on the veracity of informants' reports collected within comprehensive assessments. Indeed, similar to adults, in child and adolescent social anxiety assessments, low cross-informant correspondence has been attributed to social desirability concerns with self-reports (e.g., DiBartolo et al., 1998). As such, instances in which self-reports reflect relatively lower levels of social anxiety symptoms relative to the reports of other informants (e.g., parent) has been interpreted as a marker of a social desirability bias (De Los Reyes et al., 2012). As another example, consider assessments of neurodevelopmental conditions such as the autism spectrum disorders, which often presume the presence of signs and symptoms of such high severity that it is assumed that clients present with concerns invariantly across contexts (e.g., home and school; APA, 2013). Here, this assumption of invariance of symptoms across contexts leads many to assume, by extension, that within-informant symptom reports taken by parents or teachers are, in and of themselves, reflective of symptoms that manifest generally or across contexts (see Gotham et al., 2012). This assumption even occurs if the informants involved in the assessment typically observe those undergoing evaluation within a specific context (e.g., parents at home, teachers at school). Questions as to the veracity of informants' reports extend to reports taken about risk and protective factors of child and adolescent mental health (e.g., parenting; Taber, 2010) as well as clinical assessments conducted within specific settings (e.g., parent and adolescent reports taken within psychiatric inpatient and foster care settings; Jones et al., 2019; McWey et al., 2018). The prevailing notion in each of these literatures has to do with the negative outcomes that might transpire should one endorse the domains being assessed.

Specifically, divergent assessment information is thought to reflect bias specific to some informants (e.g., adolescents in psychiatric inpatient settings and parents in foster care settings) who might under-report concerns so as to avoid deleterious outcomes (e.g., admission into an inpatient unit, loss of child custody). Key to interpretations of informant biases in social anxiety, neurodevelopmental, psychiatric inpatient, and foster care assessments, is that many make these interpretations without carrying out direct empirical tests of these interpretations. Indeed, these interpretations are often challenging to directly test but appear face valid and are thus intuitive to researchers and clinicians. Consequently, we hold that many accept these interpretations as true, even without empirical evidence to support the interpretation.

Taken together, work examining divergent outcomes in comprehensive assessments of child, adolescent, and adult mental health often focuses on identifying methodological issues that call into question the veracity of one or more of the informants' reports taken in the assessment. In fact, in cases in which this represents a sound evidence-based interpretation of low cross-informant correspondence, one could take a variety of readily available analytic approaches to reconciling the low correspondence. For instance, if measurement error or rater bias accounts for divergent outcomes, the *unique variance* attributable to each informant could essentially be treated as nuisance variance within an integrative analytic approach, such as structural equations modeling, composite scoring, or a combinational algorithm (e.g., AND/OR rule; for reviews, see De Los Reyes et al., 2013a, 2015). Importantly, measurement error and/or rater biases represent entirely plausible *hypotheses* about divergent assessment outcomes. That is, these are merely hypotheses, not *universal truths*. These are also *testable* hypotheses, and like any hypothesis, one should not assume that they reflect reality in the absence of supportive evidence. As we note in Table 1, a key opportunity to advance the literature on comprehensive assessments of adult

mental health lies in the availability of conceptual and measurement models developed in other literatures. These models have facilitated the development of research paradigms for not only testing what divergent assessment outcomes might reflect, but also for integrating multi-informant data to optimize prediction of clinical indices. In turn, the models we describe below allow one to start treating divergent assessment outcomes not merely as methodological nuisances but rather as tools for enhancing our understanding of clients' clinical presentations.

Conceptual and Measurement Models of Comprehensive Mental Health Assessments

Throughout the history of comprehensive assessments and across many areas of inquiry in Psychology, one finds instances in which informants and methods are "pitted" against one another, what we have termed "methodological boundary disputes" (for a review, see De Los Reyes et al., 2019a). Perhaps the best example comes from research on self-reports of mental health. As previously mentioned, based on established criteria, many self-report measures exhibit sound properties in their own right (e.g., Hunsley & Mash, 2008, 2018). Interestingly, these evaluative criteria are identical to those for evaluating psychometric instruments more generally (e.g., internal consistency, construct validity, convergent and divergent validity; see Nunnally & Bernstein, 1994). Yet, this research literature exists alongside another literature: one that casts doubt on the veracity of data from self-reports, although not via well-accepted standards for evaluating the psychometric properties of instruments. Rather, the standards used appear principally focused on noting that self-reports display relatively low correspondence with estimates taken from independent assessment modalities (e.g., performance-based tasks, observed behavior; for reviews, see Dunning et al., 2004; Paulhus & Vazire, 2007). In these cases, it is as if researchers have divorced the evidence supporting the psychometric properties of self-report instruments from the very *people* completing the instruments.

The research and theory that we review below seeks to reconcile these issues with the relatively low correspondence among measures administered within comprehensive assessments, and by posing a question: What if the dispute lies not with the informants or methods one uses, but rather with the meaning (or lack thereof) that underlies divergent assessment outcomes? Indeed, consider instances in which the content coverage of self-report measures and/or those of other informants (e.g., spouses, parents, teachers) differs from the content coverage of observational or performance-based tasks, such that tasks assess behaviors displayed within relatively circumscribed or specific contexts (e.g., home, workplace, peer interactions; see also Cannon et al., 2020; MacLeod et al., 2019; Mischel & Shoda, 1995). In these instances, one can increase the interpretability of low correspondence among measures by capitalizing on the circumscribed nature of putatively "objective" measures like estimates taken from observed behavior, official records, or performance-based tasks. This is because within a multi-informant assessment battery, any one informant also tends to make their reports based on a specific observational context and clients' functioning often varies across these contexts (e.g., coworker reports on clients' behaviors in the workplace). Achenbach and colleagues (1987) termed this phenomenon "situational specificity" (see also Achenbach et al., 2005; Kraemer et al, 2003).

Specifically, rather than pitting informants and methods against one another, the work we review below identifies links among patterns of informants' subjective reports and estimates taken from non-subjective methods (e.g., observed behavior, official records, performance-based tasks). We start with a conceptual model that allows one to form testable hypotheses about the meaning underlying patterns of convergence and divergence among information sources. We then discuss a measurement model for selecting and integrating information sources to enhance the clinical utility of comprehensive mental health assessments. Throughout, we emphasize the empirical support of the models, which

largely derive outside the adult personality and psychopathology literature. Nonetheless, we conclude our chapter by highlighting ways in which these models may advance research on comprehensive assessments of adult mental health.

Operations Triad Model

A key first step in extending research on comprehensive mental health assessments involves testing hypotheses about the meaning underlying patterns of convergence and divergence among information sources. Key to this process is empirically testing whether the evidence indicates that divergence (and convergence) among information sources reflects meaningful information (e.g., captures important aspects of a client's clinical presentation) to a greater degree than it does non-meaningful information (e.g., methodological artifacts, informant biases). The Operations Triad Model (OTM; De Los Reyes et al., 2013a) offers a framework to accomplish this task, which we graphically depict in Figure 1. Specifically, the OTM allows one to form a priori hypotheses about whether informants will converge or diverge in their reports, as well as hypotheses about the reasons for convergence or divergence between their reports. In this way, the OTM offers a "gating procedure" for determining whether collecting and interpreting multi-informant assessments optimizes the clinical utility of comprehensive assessment data.

Importantly, assessment conditions vary widely by the domain(s) assessed, informants used, and settings. The likelihood is quite low that all divergent assessment outcomes stem from a single factor or set of factors. As such, embedded in the OTM are several measurement conditions that allow users of the model to hypothesize about the meaning underlying patterns among information sources. First, *Compensating Operations* reflect instances when patterns of inconsistency across information sources reflect methodological features of the assessments used. For example, a client and collateral informant's reports may differ when they complete two

distinct personality measures. Indeed, correlations between self- and other-reports are approximately .45 when informants complete the same measure and approximately .30 when informants complete unique measures (Achenbach et al., 2005). Similarly, in the child and adolescent literature, one finds significantly larger levels of cross-informant correspondence when informants complete dimensional versus discrete measures, and this is true for both internalizing (*r*s = .29 vs. .06) and externalizing (*r*s = .37 vs. .06) assessments (De Los Reyes et al., 2015).

Divergent outcomes that reflect Compensating Operations occur for several reasons, including unique item content, reliability, scaling, and scoring, even when measures purportedly assess the same construct. Thus, these differences would not reflect meaningful information that could be used to enhance our understanding of clients' clinical presentations. Rather, support for Compensating Operations would suggest that minimizing or erasing divergent outcomes among information sources would lead to a more psychometrically sound assessment. As reviewed previously, research over the past several decades has focused almost exclusively on Compensating Operations hypotheses, usually with the assumption that divergent outcomes among information sources reflect informant bias or measurement error (De Los Reyes, 2011). Thus, the OTM specifies conditions for exactly these kinds of hypotheses to provide a guide for empirically testing this predominant assumption of why divergent outcomes occur.

The other two OTM measurement conditions reflect instances when patterns of convergence and divergence reflect meaningful information (De Los Reyes et al., 2013a). *Converging Operations* reflects a set of measurement conditions for interpreting patterns of consistent information sources as accurate reflections of the same conclusion or meaningful clinical phenomena. For example, a client and their family member's reports may converge in

high levels of personality symptoms if these two informants' reports capture behavior in the same context (e.g., home) or reflect a more severe and impairing symptom presentation. As another example, consider a persistent form of psychopathology that manifests across disparate contexts (e.g., schizophrenia). Here, a Converging Operations hypothesis might involve expecting two informants who observe behavior in different contexts (e.g., client's spouse and coworker) to provide similar reports about the presence of psychosis symptoms.

Alternatively, *Diverging Operations* reflects a set of measurement conditions for interpreting patterns of inconsistent information sources based on hypotheses about variations in the behavior being assessed. For example, a client's coworker and family member may provide reports of personality symptoms that diverge from one another given that they each observe the client's behavior in unique contexts (e.g., home, work). Further, these reports may diverge if the conditions across contexts differ to such an extent that the client behaves differently within each context. Alternatively, a client may report far fewer personality symptoms than their romantic partner when, consistent with the client's symptom presentation, they have low insight on others' perceptions of their own behavior. In this instance of Diverging Operations, the divergence between the client and their romantic partner represents meaningful clinical information that can inform clinical decision-making. When support for Converging and Diverging Operations is found, one would want to avoid minimizing or erasing differences among information sources when making clinical decisions. Instead, leveraging information about the meaning underlying patterns among information sources would enhance the utility of information obtained in a comprehensive mental health assessment.

As summarized in Table 1, a burgeoning area of research drawing on the OTM and largely in research on child and adolescent mental health directly tests the veracity of divergent

outcomes among multi-informant reports. Much of this research focuses on individual differences and uses person-centered approaches to examine how patterns of reports relate to clinically relevant indices. Such approaches are useful in OTM-informed research and reveal that, although informants exhibit overall low-to-moderate levels of convergence in their reports, they display within-sample individual differences in patterns among their reports. Specifically, multi-informant assessments consistently yield patterns between informants' reports characterized by (a) convergence in low or high levels of problems or (b) divergence such that one informant reports higher levels of problems than the other (Makol, De Los Reyes, et al., 2020). Examining these patterns as they relate to contextual information yields support for Converging and Diverging Operation hypotheses when interpreting multi-informant reports. For example, De Los Reyes and colleagues (2009) found that parent and teacher reports of preschoolers' disruptive behavior relates to children's disruptive behavior observed in home and school contexts. That is, supporting that informants' reports meaningfully reflect the context in which they observe youth, the children in this study tended to display disruptive behavior in contexts in which informants reported elevations in disruptive behavior. In contrast, children tended to *not* display disruptive behavior in contexts in which informants did not report elevations in disruptive behavior. Further, when rating children's social functioning, parents, teachers, and peers tend to provide reports that relate to context-relevant social functioning (Hartley et al., 2011; Kwon et al., 2012).

Research on multi-informant reports of social anxiety reveals similar findings. For example, adolescents and unfamiliar peers, and not parents, provide reports of adolescent social anxiety that predict adolescent arousal in social interactions with unfamiliar peers (Deros et al., 2018). Further, convergence in clinician and adult client reports in high levels of social anxiety

symptoms are associated with an increased likelihood that clients display social skills deficits across social interaction contexts (De Los Reyes et al., 2013b). Thus, a growing evidence-base supports the possibility that variations among informants' reports reflect two key features of behaviors assessed in comprehensive mental health assessments: (1) behaviors vary across contexts and (2) informants vary in their opportunities to observe behaviors across the contexts in which they observe individuals undergoing evaluation. By extension, when the evidence supports interpreting divergent assessment outcomes as reflections of contextual variations in behavior (i.e., a Diverging Operations hypothesis), treating differences among informants' reports as error or bias would only reduce one's ability to accurately characterize clients' behavior.

Beyond context, examining variations among informants yields other meaningful information about clients' clinical presentations. For example, patterns among parents' and adolescents' reports of mood and suicidal ideation and behavior relate to the likelihood that youth receive mental health services (Jones et al., 2019; Makol & Polo, 2018). Patterns among parent and teacher reports of youth autism spectrum disorder symptoms relate to clinical severity indices including medication use and diagnosis (Lerner et al., 2017). Interestingly, divergence in informants' perspectives portends risk for poorer outcomes. For example, when parents report higher knowledge of their adolescents' whereabouts and activities than adolescents report about their parents' knowledge, adolescents are at higher risk for later developing substance use problems (Lippold et al., 2013, 2014).

Other research supports that patterns among informants' reports at intake relate to clients' treatment course. For instance, youth who self-report lower levels of anxiety symptoms compared to their parents at intake tend to experience poorer treatment outcomes (i.e., relative to

other patterns of parent-youth intake reports; Becker-Haimes et al., 2018). These findings suggest that divergence between informants' reports may serve as a marker for relatively low engagement of youth clients in treatment. Similarly, in psychiatric inpatient settings, patterns of parent and adolescent reports of internalizing problems at intake relate to adolescent suicide status, diagnosis, and treatment outcomes over the course of adolescents' inpatient care (Makol et al., 2019). In this study, adolescents reporting lower levels of symptoms relative to their parents were more likely to receive standing antipsychotics and seclusion during their inpatient stay. These findings demonstrate that examining patterns among informants' reports at intake can provide crucial information that informs prognosis and treatment planning.

Overall, prior work on the OTM indicates that examining patterns of convergence and divergence between multiple informants' reports provides incrementally valuable information. By definition, one cannot obtain this information when conducting assessments based on one informant's report or when conducting comprehensive assessments that result in disregarding or ignoring divergent outcomes. Additional support for the central tenets of, and empirical support for, the OTM comes from recent work that extends the OTM to literatures that involve interpreting (a) peripheral physiological data in clinical assessments (De Los Reyes & Aldao, 2015), (b) multi-informant assessments of family functioning (De Los Reyes & Ohannessian, 2016; De Los Reyes et al., 2019b), (c) school-based assessments of psychosocial functioning (De Los Reyes et al., 2019c) and (d) strategies for integrating multivariate biobehavioral data collected as part of the National Institute of Mental Health's Research Domain Criteria (De Los Reyes et al., 2020).

Kraemer and Colleagues' "Satellite" Model

After finding evidence that information sources in comprehensive mental health assessments individually and collectively provide valid information about the clients undergoing evaluation, clinicians and researchers are still tasked with determining how to validly integrate and interpret these reports. As noted in Table 1, there is a lack of consensus on how to best integrate information in comprehensive assessments of adult personality and psychopathology. As previously described, integrative approaches vary such that they each draw upon unique assumptions about the meaning underlying divergent assessment outcomes. Thus, users of integrative approaches designed to "deal with" these divergent outcomes should carefully scrutinize the underlying assumptions of these approaches. That is, when selecting an approach for integrating multi-informant data, it is incumbent upon users to ensure that any approach used is consistent with empirical work on what divergent assessment outcomes reflect. Linking the integrative approach to both its assumptions regarding divergent outcomes and the evidence supporting these assumptions optimizes the clinical utility of information obtained from these assessments.

Kraemer and colleagues' (2003) "Satellite" Model offers an integrative strategy that aims to optimize one's use of multiple information sources. As with a Diverging Operations hypothesis within the OTM, the Satellite Model assumes that differences among informants' reports reflect meaningful information (see Figure 2). Along these lines, the Satellite Model draws on principles from the geographic sciences, specifically the global positioning system (GPS), in proposing strategies for both selecting and integrating information sources. Specifically, when identifying a target location in space (e.g., a building), one must use satellites that systematically vary in their latitude and longitude. One does not precisely locate a target using a single satellite, given that satellites rarely hover over the exact location being targeted. Similarly, one also does not precisely locate a target by placing multiple

satellites at a single latitude/longitude position. Instead, one effectively *triangulates* on a target's position by obtaining reliable and valid information from three satellites that vary in their latitude and longitude, relative to the target.

When applying these concepts to comprehensive mental health assessments, Kraemer and colleagues (2003) argue that one effectively captures a target's mental health functioning through selection of information sources that systematically vary in key criteria, what they refer to as a "mix-and-match" criterion. The criterion they lay out is the metaphorical equivalent to the latitude and longitude parameters used to strategically position satellites to locate an object in geographic space. For example, one could effectively triangulate on a target's mental health by selecting informants who systematically vary in the context (e.g., home, school) and perspective (e.g., self, other) from which they rate youth mental health (see Figure 2a). Kraemer and colleagues argue that one should select information sources that they *expect* to disagree, in an effort to enhance the clinical utility of an assessment. Stated otherwise, rather than considering divergence among information sources as a *barrier* against sound decision-making, divergence becomes a *tool* that facilitates sound decision-making. This represents an innovative departure from historical thinking on divergent assessment outcomes, namely that they reflect measurement error or informant bias (De Los Reyes, 2011).

To implement the Satellite Model, Kraemer and colleagues (2003) apply principal components analysis (PCA) to a strategically selected set of informants' reports (see Figure 2b). In doing so, they demonstrated that informants' reports can be decomposed into three systematic and meaningful components. First, a *Trait* component captures concerns that manifest across contexts and perspectives. Second, a *Context* component captures the environment in which the informants observe the target (e.g., home vs. non-home). Third, a *Perspective* component captures the vantage point from which the informants observe the target (e.g., self vs. observer). Although one leverages PCA to identify three

components, Kraemer and colleagues designed the Trait score to have maximal utility in estimating or predicting clinical outcomes.

Makol, Youngstrom, et al. (2020) recently applied the Satellite Model to multi-informant assessments of adolescent social anxiety (Figure 3). In their application of the approach, Makol, Youngstrom, et al. selected three informants who varied in the contexts and perspectives from which they observe adolescent social anxiety symptoms. This included parents and adolescents, who represent commonly utilized informants in social anxiety assessments, as well as unfamiliar peer confederates who interacted with adolescents in a controlled set of social interaction tasks designed to elicit anxiety within social interactions with unfamiliar peers (for a detailed description, see Cannon et al., 2020). These informants were selected given that they met Kraemer et al.'s (2003) "mix-and-match" criterion. Specifically, parents and peer confederates each perceive the adolescent from an observer perspective but from unique contexts (i.e., home vs. peer), while adolescents perceive their behavior from a self-perspective and across home and peer contexts. Makol, Youngstrom, et al. found that the Trait score optimized prediction of clinical variables relevant to understanding adolescent social anxiety. Specifically, the Trait score outperformed individual informants' reports and a composite score of informants' reports in predicting observed anxiety in social interaction tasks (βs = .47-.67) and referral status (i.e., whether adolescents were part of a clinic-referred or community control group; ORs = 2.66-6.53), with effects in the large-magnitude range. Although Kraemer et al.'s Satellite Model is a historically underutilized approach; Makol, Youngstrom, et al. demonstrate its potential value as an integrative strategy within comprehensive mental health assessments.

Applying the OTM and Satellite Models to Personality and Psychopathology Assessment

In this chapter, we reviewed research on divergent outcomes observed in comprehensive mental health assessments. In particular, we leveraged the deep literature on divergent outcomes as they

manifest in comprehensive assessments of child and adolescent mental health to highlight ways in which we might improve our understanding of divergent outcomes in assessments of adult personality and psychopathology. Specifically, research findings in child and adolescent assessment gave rise to a conceptual model (i.e., OTM; De Los Reyes et al., 2013a) designed to facilitate building an evidence base for validly interpreting patterns of divergent and convergent assessment outcomes. Applying the OTM to adult personality and psychopathology assessments will involve reconsidering not only how to interpret divergent outcomes but also the process of administering these assessments. In particular, drawing inferences as to the meaning underlying divergent outcomes between two informant's reports will involve comparing patterns of these reports with data derived from other sources, such as clients' criminal records in forensic assessments or data derived from performance-based tests in neuropsychological assessments. Beyond incorporating data from independent sources, we encourage careful consideration of the contexts of observation to which informants have access. When informants used in a comprehensive assessment vary as to their contexts of observation, it is critical to choose independent sources that one can leverage to test hypotheses about what patterns of convergence or divergence between informants' reports might represent.

As with practitioners and researchers focused on assessing child and adolescent mental health, those focused on assessing personality and psychopathology often harbor assumptions about the reasons that informants disagree. Historically, these assumptions have focused on the ways in which disagreement signals bias or measurement error. As an example, consider the assumption that low convergence between self- and other-reports of personality traits is a marker of low insight on the part of the client being rated (Carlson & Oltmanns, 2015; Samuel et al., 2018). One application of the OTM to comprehensive assessments of adult mental health could involve testing this assumption. This

may involve clarifying what the link between "low insight" and relatively low levels of correspondence between self- and other-reports reflects.

Traditionally, "inaccurate" self-reports associated with low insight would be conceptualized within a Compensating Operations hypothesis. However, might a large divergence between self- and other-reports reflect clinically meaningful information, even in the case of low insight on the part of a specific informant's report (e.g., client)? In particular, might this divergence serve as a marker for the severity, type, and impairment associated with personality pathology (e.g., Fischer et al., 2020)? Importantly, consider that in classical test theory, the presence of bias directly impacts measurement validity or accuracy in measurement, insofar as a factor (e.g., low insight) systematically reflects characteristics that are superfluous to the purpose of the assessment (e.g., understanding and characterizing a client's concerns; Nunnally & Bernstein, 1994). Thus, an OTM-informed study addressing these questions would need to distinguish an informant's reporting "biases" that are unrelated to clients' clinical presentations from those that are related to their clinical presentations. To do so, one might examine how patterns of informants' reports of personality pathology relate to independent sources relevant to the condition being assessed. For instance, when collecting data from collateral informants who observe the client in different contexts (e.g., spouse vs. coworker), researchers may select independent sources focused on capturing information about behavioral displays in the home (e.g., couples interaction task) and workplace (e.g., hours worked, customer complaints). As another example, researchers may collect data on treatment outcomes (e.g., dropout rates, diagnoses at discharge, psychiatric medication adherence) to determine whether patterns among self- and other-reports at intake inform prognosis and the course of treatment. We encourage work that rigorously tests hypotheses about the meaning underlying multiple information sources in comprehensive mental health assessments. As in the child and adolescent

mental health field, we surmise that patterns among informants will yield clinically meaningful information. In the case of personality psychopathology, differences in perceptions of a target's behavior may prove clinically useful and inform clinical decision-making (Ashton et al., 2017; McCrae, 2018).

Research findings in child and adolescent assessment also inform measurement models designed to integrate information sources whose reports diverge from each other in predictable ways. For instance, consider steps one might take following a line of OTM-inspired research that reveals measurable factors that one can use to predict divergent outcomes between reports about clients' aggressive behavior completed by spouses and coworkers. Like in the example used to highlight integrative models for child and adolescent mental health (Figure 2), the idea here would be to identify the "latitudes and longitudes" that facilitate triangulation among the informants used in the assessment. That is, what factors predict whether informants used in the assessment will provide divergent outcomes? In this case, the spouse's and coworker's reports might disagree with each other in large part due to observing the client at home and work, respectively. Yet, these two informants have one thing in common: both provide reports from the perspective of an observer of the client. Adding an informant whose report can be conceptualized as a "mix" of observations of the client in the home and work context, but from a perspective that differs from that of spouse and coworker (e.g., client self-report), would allow an assessor to arrive at a triangulated estimate of the client's levels of aggressive behavior (see Figure 3). Identifying how to select information sources that systematically vary and meet the "mix-and-match" criterion will likely differ depending on the personality domain being assessed. For example, informants selected in an assessment of borderline personality disorder may meaningfully and systematically differ in their perspective (e.g., self vs. other) and relationship to the client (e.g. romantic vs. familial). As another example, consider that information sources selected in an assessment of

antisocial personality disorder may differ in the time at which they were collected (e.g., adolescence vs. adulthood) and context (e.g., home vs. work). When comparing both of these cases, the "latitudes and longitudes" that inform selection of informants might vary. Yet, their shared goal involves identifying informants who provide reports that, when integrated, triangulate on the domain being assessed.

We began our chapter with an example of a common assessment scenario. We used this example to highlight the interpretive issues that arise with current strategies for collecting and understanding comprehensive mental health assessment data. Our overview of an emerging body of evidence, frameworks, and measurement models that facilitate interpretation and integration of comprehensive assessment data reveals concrete strategies for reimagining our assessment procedures. We illustrate these innovative procedures with a case example that applies the conceptual and measurement models discussed previously to a comprehensive mental health assessment of schizophrenia. Consider that you are a clinician in a hospital who is conducting an intake assessment for Alex, a 24-year-old male who graduated from college two years prior and currently works as a software developer. Alex was referred to you given a suspected first-episode psychosis. Alex recently received negative performance reviews by his supervisor and, subsequent to these reviews, he took a leave of absence from work. For this reason, he is at risk of losing his job. Alex's parents first became concerned about his behavior a few months prior. Specifically, when visiting his apartment and observing unsanitary conditions, they also learned that Alex spent no time socializing with others. Following this visit, Alex moved back in with his parents. When interacting with Alex in recent months, his mother noticed apparent delusions that others were "out to get him" at work. They also noticed Alex's inability to complete basic hygiene tasks as well as displays of unusual behaviors and disorganized speech. In addition to arriving at a diagnostic formulation to characterize Alex's challenges, you must also determine whether Alex would benefit from placement in one or more

psychosocial intervention programs offered through the hospital. Social skills deficits are a key associated feature of schizophrenia (Meyer & Kurtz, 2009). Thus, you seek to leverage knowledge about social skills deficits and their links to schizophrenia to not only determine whether a social skills intervention is appropriate for Alex's presenting concerns, but also to make sense of the multi-informant assessments you administer to characterize Alex's concerns.

How might a clinician design an assessment battery aimed at characterizing Alex's social functioning? One option might be to leverage the Satellite Model to identify informants who, collectively, triangulate on a sound measure of Alex's social functioning. Specifically, the clinician might administer a multi-informant assessment battery that systematically varies the perspectives (e.g., self vs. other) and contexts (e.g., home vs. work) from which the informants provide reports about Alex's social functioning. These measures might include the Social Functioning Scale (SFS; Birchwood et al., 1990), a self-report instrument that Alex completes to provide information from a self-perspective about home and work functioning. These measures might also include the Social Behavior Schedule (SBS; Wykers & Sturt, 1986), a collateral report instrument that (a) Alex's mother completes to provide information from an other-perspective about Alex's home functioning, and (b) Alex's coworker completes to provide information from an other-perspective about Alex's work functioning.

Imagine that after collecting these reports, the clinician finds that Alex's mother and coworker report greater and clinically significant levels of social impairment, relative to Alex's self-report. How might the clinician interpret these discrepancies in a way that facilitates sound clinical decision-making? Levels of insight vary across individuals with schizophrenia, and this variation predicts unique illness trajectories (Murri et al., 2015). With this knowledge, the clinician could use divergent outcomes in reports of social functioning to conceptualize Alex's

level of insight regarding his social functioning. One method for testing this hypothesis might be to administer a performance-based measure to Alex to assess social functioning, such as the Metacognitive Assessment Scale (MAS; Semerari et al., 2003). Consistent with prior work (Fischer et al., 2020), the clinician poses a Diverging Operations hypothesis: Discrepancies between Alex's self-reports of social functioning and those of other informants serves as a marker for low insight, a hypothesis supported by Alex's relatively low scores on the MAS. The clinician may recommend Alex for a social skills intervention, but first work with him on psychoeducation to both raise his awareness about his social difficulties and increase his motivation for participating in the intervention.

As an alternative outcome of this assessment, imagine that the clinician finds that Alex self-reports a fair degree of social impairments, his coworker reports relatively high levels of social impairments, and his mother reports relatively low levels of social impairments. How might the clinician interpret these discrepancies? One hypothesis may be that the environmental demands placed on Alex at home substantially differ from those placed on him at work. The clinician may consider administering the Assessment of Interpersonal Problem-Solving Skills (AIPSS; Donahue et al., 1990) to Alex to assess social problem solving skills that may relate more to his work functioning than his home functioning (e.g., problem solving during performance-based scenarios). The clinician might also administer a family discussion task in order to detect maladaptive interactions among family members (e.g., Kiecolt-Glaser et al., 1987; Thomas et al., 2017). Consistent with prior work on social anxiety assessments (De Los Reyes et al., 2013b), the clinician poses a Diverging Operations hypothesis: Discrepancies between parent and coworker reports of social functioning map onto Alex's poor performance on the AIPSS but relatively normative performance on the family discussion task. Specifically, the task data

reveals that Alex's performance reflects demands placed on him at work (e.g., collaborating with coworkers to complete a project), but not at home (e.g., completing chores with high levels of parent support). The clinician may recommend Alex for a social skills intervention embedded within a work rehabilitation program to support the development of social skills and increase the likelihood that he can continue his work as a software developer in the future.

This case example illustrates how one can leverage the conceptual and measurement models described in this chapter to revise procedures for administering and interpreting comprehensive mental health assessment data, and in a way that facilitates use of assessment outcomes to guide mental health service delivery. Indeed, many assessment batteries could be redesigned to assess domains of psychopathology and personality, and specific assessments used will vary based on the questions at hand. We encourage the reader to practice applying these models to commonly assessed domains in comprehensive assessments of adult mental health.

Concluding Comments

In our review of work on divergent assessment outcomes in comprehensive mental health assessments, we have discussed these matters with regard to both practice and research. We advanced conceptual and measurement models designed to improve the interpretability of comprehensive assessments administered in various settings. Indeed, we previously discussed that practitioners and researchers alike have access to decades of studies on the levels and moderators of cross-informant correspondence (Achenbach et al., 1987, 2005; De Los Reyes et al., 2015, 2019a). Importantly, these effects manifest regardless of the setting in which one administers these assessments, from controlled laboratory work to work carried out in a host of service settings (e.g., hospitals, medical centers, community mental health clinics). As such, these data allow one to come to informed predictions about whether a comprehensive assessment will reveal divergent outcomes and, if so, what these outcomes

might mean. In these respects, we encourage modifications to administrations of assessments in both practice and research settings. In fact, assessors across settings might dramatically improve use and interpretation of data from comprehensive mental health assessments by restructuring assessment procedures, and well in advance of administering these assessments.

Specifically, all assessors should hypothesize what the outcomes of comprehensive assessments might reveal. In doing so, gathering information about factors hypothesized to drive divergent outcomes may prove instrumental when integrating and interpreting any data collected. Our case example provides a window into reimagining the assessment process, and it begins at the point of the referral and intake process. As assessors, we should be learning, in advance of the assessment, of the contingencies that elicit clients' needs for services (e.g., interpersonal conflict). We should seek to understand if these contingencies manifest invariantly across contexts (e.g., in the workplace and home) or to varying degrees across contexts (e.g., home > workplace). Learning this information in advance of administering measures to multiple informants (e.g., client, coworker, spouse) will have the power of reducing uncertainties in interpreting cross-informant reports, particularly if one informant (coworker) but not the other (spouse) endorses the presence of the behaviors being assessed. Further, reducing these uncertainties with characterizing clients' concerns will also facilitate treatment planning and monitoring treatment response, particularly if you learn that treatment should focus on a specific context (e.g., workplace), with continued monitoring in contexts where the target behaviors do not currently manifest (e.g., home). In sum, we encourage practitioners and researchers alike to leverage assessment strategies designed to decrease their uncertainty with how to make clinical decisions when comprehensive assessments yield divergent outcomes.

References

Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science, 15*(2)*,* 94-98.

Achenbach, T. M. (2017). Future directions for clinical research, services, and training: Evidence-based assessment across informants, cultures, and dimensional hierarchies. *Journal of Clinical Child and Adolescent Psychology, 46*(1)*,* 159-169.

Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin, 131*(3)*,* 361-382.

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*(2), 213-232.

Alexander, L. A., McKnight, P. E., Disabato, D. J., & Kashdan, T. B. (2017). When and how to use multiple informants to improve clinical assessments. *Journal of Psychopathology and Behavioral Assessment*, *39*(4), 669-679.

Al Ghriwati, N., Winter, M. A., Greenlee, J. L., & Thompson, E. L. (2018). Discrepancies between parent and self-reports of adolescent psychosocial symptoms: Associations with family conflict and asthma outcomes. *Journal of Family Psychology*, *32*(7), 992-997.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

Ashar, Y. K., Chang, L. J., & Wager, T. D. (2017). Brain mechanisms of the placebo effect: An affective appraisal account. *Annual Review of Clinical Psychology*, *13*, 73-98.

Ashton, M. C., De Vries, R. E., & Lee, K. (2017). Trait variance and response style variance in the scales of the personality inventory for DSM–5 (PID–5). *Journal of Personality Assessment*, *99*(2), 192-203.

Balsis, S., Cooper, L. D., & Oltmanns, T. F. (2015). Are informant reports of personality more internally consistent than self reports of personality. *Assessment, 22*(4), 399-404.

Becker-Haimes, E. M., Jensen-Doss, A., Birmaher, B., Kendall, P. C., & Ginsburg, G. S. (2018). Parent–youth informant disagreement: Implications for youth anxiety treatment. *Clinical Child Psychology and Psychiatry*, *23*(1), 42-56.

Birchwood, M., Smith, J., Cochrane, R., Wetton, S., & Copestake, S. (1990). The Social Functioning Scale: The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *The British Journal of Psychiatry*, *157*, 853-859.

Birchwood, M., Smith, J., Drury, V., Healy, J., Macmillan, F., & Slade, M. (1994). A self-report Insight Scale for psychosis: Reliability, validity and sensitivity to change. *Acta Psychiatrica Scandinavica*, *89*(1), 62-67.

Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: Questions and answers for the DSM-V. *Depression and Anxiety, 27*(2), 168-189.

Carlson, E. N., & Oltmanns, T. F. (2015). The role of metaperception in personality disorders: Do people with personality problems know how others experience their personality? *Journal of Personality Disorders, 29*(4), 449-467.

Cannon, C. J., Makol, B. A., Keeley, L. M., Qasmieh, N., Okuno, H., Racz, S.J., & De Los Reyes, A. (2020). A paradigm for understanding adolescent social anxiety with

unfamiliar peers: Conceptual foundations and directions for future research. *Clinical Child and Family Psychology Review*, *23*(3), 338-364.

Clark, L. A., Livesley, W. J., & Morey, L. (1997). Personality disorder assessment: The challenge of construct validity. *Journal of Personality Disorders*, *11*(3), 205-231.

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), 1-9.

De Los Reyes, A., & Aldao, A. (2015). Introduction to the special issue: Toward implementing physiological measures in clinical child and adolescent assessments. *Journal of Clinical Child and Adolescent Psychology, 44*(2)*,* 221-237.

De Los Reyes, A., Aldao, A., Thomas, S. A., Daruwala, S., Swan, A. J., Van Wie, M., Goepel K. A., & Lechner, W. V. (2012). Adolescent self-reports of social anxiety: Can they disagree with objective psychophysiological measures and still be valid? *Journal of Psychopathology and Behavioral Assessment*, *34*(3), 308-322.

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*(4), 858-900.

De Los Reyes, A., Bunnell, B. E., & Beidel, D. C. (2013b). Informant discrepancies in adult social anxiety disorder assessments: Links with contextual variations in observed behavior. *Journal of Abnormal Psychology, 122*(2)*,* 376-386.

De Los Reyes, A., Cook, C. R., Gresham, F. M., Makol, B. A., & Wang, M. (2019c). Informant discrepancies in assessments of psychosocial functioning in school-based services and research: Review and directions for future research. *Journal of School Psychology*, *74*,

74-89.

De Los Reyes, A., Drabick, D. A. G., Makol, B. A., & Jakubovic, R. (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*(3), 279-296.

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*(5), 637-652.

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 and Adolescent Psychology, 47*(3), 357-365.

De Los Reyes, A., Lerner, M. D., Keeley, L. M., Weber, R., Drabick, D. A. G., Rabinowitz, J., & Goodman, K. L. (2019a). Improving interpretability of subjective assessments about psychological phenomena: A review and cross-cultural meta-analysis. *Review of General Psychology*, *23*(3), 293-319.

De Los Reyes, A., & Ohannessian, C. M. (2016). Introduction to the special issue: Discrepancies in adolescent-parent perceptions of the family and adolescent adjustment. *Journal of Youth and Adolescence, 45*(10), 1957-1972.

De Los Reyes, A., Ohannessian, C. M., & Racz, S. J. (2019b). Discrepancies between adolescent and parent reports about family relationships. *Child Development Perspectives*, *13*(1), 53-58.

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013a). Principles

underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology, 9*, 123-149.

Deros, D. E., Racz, S. J., Lipton, M. F., Augenstein, T. M., Karp, J. N., Keeley, L. M., Qasmieh, N., Grewe, B. I., Aldao, A., De Los Reyes, A. (2018). Multi-informant assessments of adolescent social anxiety: Adding clarity by leveraging reports from unfamiliar peer confederates. *Behavior Therapy, 49*(1), 84-98.

DiBartolo, P. M., Albano, A. M., Barlow, D. H., & Heimberg, R. G. (1998). Cross-informant agreement in the assessment of social phobia in youth. *Journal of Abnormal Child Psychology, 26*(3)*,* 213-220.

Donahoe, C. P., Carter, M. J., Bloem, W. D., Hirsch, G. L., Laasi, N., & Wallace, C. J. (1990). Assessment of interpersonal problem-solving skills. *Psychiatry*, *53*(4), 329-339.

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice, 7*(4)*,* 435-453.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69-106.

Fischer, M. W., Dimaggio, G., Hochheiser, J., Vohs, J. L., Phalen, P., & Lysaker, P. H. (2020). Metacognitive capacity is related to self-reported social functioning and may moderate the effects of symptoms on interpersonal behavior. *The Journal of Nervous and Mental Disease*, *208*(2), 138-142.

Galione, J. N., & Oltmanns, T. F. (2013). Identifying personality pathology associated with major depressive episodes: Incremental validity of informant reports. *Journal of Personality Assessment, 95*(6)*,* 625-632.

Ganellen, R. J. (2007). Assessing normal and abnormal personality functioning: Strengths and weaknesses of self-report, observer, and performance-based methods. *Journal of Personality Assessment*, *89*(1), 30-40.

Gotham, K., Pickles, A., & Lord, C. (2012). Trajectories of autism severity in children using standardized ADOS scores. *Pediatrics*, *130*(5), 1278-1284.

Granholm, E. L., McQuaid, J. R., & Holden, J. L. (2016). *Cognitive-behavioral social skills training for schizophrenia: A practical treatment guide*. Guilford Publications.

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 and Adolescent Psychology*, *40*(1), 54-66.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*(1), 78-79.

Hou, Y., Benner, A. D., Kim, S. Y., Chen, S., Spitz, S., Shi, Y., & Beretvas, T. (2019). Discordance in parents' and adolescents' reports of parenting: A meta-analysis and qualitative review. *American Psychologist, 75*(3), 329-348.

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3*(1), 29-51.

Hunsley, J., & Mash, E. J. (Eds.). (2018). *A guide to assessments that work* (2nd ed.). New York, NY: Oxford University Press.

Hunsley, J., & Mash, E. J. (Eds.). (2008). *A guide to assessments that work*. New York, NY: Oxford University Press.

Huprich, S. K., Bornstein, R. F., & Schmitt, T. A. (2011). Self-report methodology is insufficient for improving the assessment and classification of Axis II personality disorders. *Journal*

of Personality Disorders, *25*(5), 557-570.

Jones, J. D., Boyd, R. C., Calkins, M. E., Ahmed, A., Moore, T. M., Barzilay, R., Benton, T. D., & Gur, R. E. (2019). Parent-adolescent agreement about adolescents' suicidal thoughts. *Pediatrics*, *143*(2), 1-12.

Kiecolt-Glaser, J. K., Fisher, L. D., Ogrocki, P., Stout, J. C., Speicher, C. E., & Glaser, R. (1987). Marital quality, marital disruption, and immune function. *Psychosomatic Medicine*, *49*(1), 13-34.

Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and Practice, 9*(3), 300-311.

Korelitz, K. E., & Garber, J. (2016). Congruence of parents' and children's perceptions of parenting: A meta-analysis. *Journal of Youth and Adolescence*, *45*(10), 1973-1995.

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*(9), 1566-1577.

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*(3), 121-133.

Lapouse, R., & Monk, M. A. (1958). An epidemiologic study of behavior characteristics in children. *American Journal of Public Health, 48*(9)*,* 1134-1144.

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*(7), 829-839.

Lippold, M. A., Greenberg, M. T., & Collins, L. M. (2013). Parental knowledge and youth risky behavior: A person oriented approach. *Journal of Youth and Adolescence*, *42*(11), 1732-1744.

Lippold, M. A., Greenberg, M. T., & Collins, L. M. (2014). Youths' substance use and changes in parental knowledge-related behaviors during middle school: A person-oriented approach. *Journal of Youth and Adolescence*, *43*(5), 729-744.

MacLeod, C., Grafton, B., & Notebaert, L. (2019). Anxiety-linked attentional bias: Is it reliable? *Annual Review of Clinical Psychology, 15,* 529-554.

Makol, B. A., De Los Reyes, A., Garrido, E., Harlaar, N., & Taussig, H. (2020). Assessing the mental health of maltreated youth with child welfare involvement using multi-informant reports. *Child Psychiatry and Human Development*. Advance online publication. https://doi.org/10.1007/s10578-020-00985-8

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*(10), 1677-1689.

Makol, B. A., & Polo, A. J. (2018). Parent-child endorsement discrepancies among youth at chronic-risk for depression. *Journal of Abnormal Child Psychology*, *46*(5), 1977-1088.

Makol, B. A., Youngstrom, E. A., Racz, S. J., Qasmieh, N., Glenn, L. E., & De Los Reyes, A. (2020). Integrating multiple informants' reports: How conceptual and measurement models may address long-standing problems in clinical decision-making. *Clinical Psychological Science*. Advance online publication. https://doi.org/10.1177/2167702620924439

McCrae, R. R. (2018). Method biases in single-source personality assessments. *Psychological Assessment*, *30*(9), 1160-1173.

McCrae, R. R., Costa Jr, P. T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., Hřebíčková, M., & Urbánek, T. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality*, *38*(2), 179-201.

McWey, L. M., Cui, M., Cooper, A. N., & Ledermann, T. (20518). Caregiver-adolescent disagreement on the mental health of youth in foster care: the moderating role of the caregiver relationship. *Child Maltreatment*, *23*(3), 294-302.

Meyer, M. B., & Kurtz, M. M. (2009). Elementary neurocognitive function, facial affect recognition and social-skills in schizophrenia. *Schizophrenia Research*, *110*(1-3), 173-179.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. *Psychological Bulletin*, *139*(3), 548-605.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*(2)*,* 246-268.

Murri, M. B., Respino, M., Innamorati, M., Cervetti, A., Calcagno, P., Pompili, M., ... & Amore, M. (2015). Is good insight associated with depression among patients with schizophrenia? Systematic Review and Meta-Analysis. *Schizophrenia Research*, *162*(1-3), 234-247.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Oltmanns, T. F., & Turkheimer, E. (2009). Person perception and personality pathology. *Current*

Directions in Psychological Science, 18(1)*,* 32-36.

Samuel, D. B., Suzuki, T., Bucher, M. A., & Griffin, S. A. (2018). The agreement between clients' and their therapists' ratings of personality disorder traits. *Journal of Consulting and Clinical Psychology, 86*(6), 546-555.

Samuel, D. B., Suzuki, T., & Griffin, S. A. (2016). Clinicians and clients disagree: Five implications for clinical science. *Journal of Abnormal Psychology, 125*(7)*,* 1001-1010.

Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review, 24*(2)*,* 239-254.

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*(3)*,* 485-499.

Roberts, B. W., & Caspi, A. (2001). Personality development and the person–situation debate: It's déjà vu all over again. *Psychological Inquiry, 12*(2), 104-109.

Romano, E., Weegar, K., Babchishin, L., & Saini, M. (2018). Cross-informant agreement on mental health outcomes in children with maltreatment histories: A systematic review. *Psychology of Violence, 8*(1), 19-30.

Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., & Alleva, G. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychology & Psychotherapy*, *10*(4), 238-261.

Stratis, E. A., & Lecavalier, L. (2015). Informant agreement for youth with autism spectrum disorder or intellectual disability: A meta-analysis. *Journal of Autism and Developmental Disorders*, *45*(4), 1026-1041.

Taber, S. M. (2010). The veridicality of children's reports of parenting: A review of factors

contributing to parent–child discrepancies. *Clinical Psychology Review, 30*(8)*,* 999-1010.

Thomas, S. A., Wilson, T., Jain, A., Deros, D. E., Um, M., Hurwitz, J., Jacobs, I., Myerberg, L., Ehrlich, K. B., Dunn, E. J., Aldao, A., Stadnik, R., De Los Reyes, A. (2017). Toward developing laboratory-based parent-adolescent conflict discussion tasks that consistently elicit adolescent conflict-related stress responses: Support from physiology and observed behavior. *Journal of Child and Family Studies*, *26*(12), 3288-3302.

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 98*(2), 281-300.

Vazire, S., & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science, 20*(2), 104-108.

Wykes, T., & Sturt, E. (1986). The measurement of social behaviour in psychiatric patients: An assessment of the reliability and validity of the SBS schedule. *The British Journal of Psychiatry*, *148*, 1-11.

Youngstrom, E., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology, 67*(6), 905-916.

Youngstrom, E. A., VanMeter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M. L., & Youngstrom, J. K. (2017). Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment. *Clinical Psychology: Science and Practice*, *24*(4), 331-363.

Table 1. Challenges and opportunities with the evidence base on interpreting comprehensive assessments of adult personality and psychopathology

Challenge	*Opportunity*	*Representative Publications*
1. Relatively nascent literature documenting levels and moderators of divergence between informants' reports of adult personality and psychopathology	1. Decades-long literature consisting of over 400 studies, documenting levels of moderators of divergence between informants' reports of child and adolescent psychopathology	• Achenbach et al. (1987) • De Los Reyes et al. (2019a) • De Los Reyes et al. (2015) • Duhig et al. (2000) • Korelitz & Garber (2016) • Lapouse & Monk (1958) • Renk & Phares (2004) • Romano et al. (2018) • Stratis & Lecavalier (2015)
2. Prevailing ideas about the relative veracity of specific informants' reports of adult personality and psychopathology	2. Evidence within assessment literatures that historically casted doubt on specific informants' reports, supporting the value of interpreting divergence between these reports and those of other informants	• De Los Reyes et al. (2013a) • Deros et al. (2018) • Makol, De Los Reyes, et al. (2020) • Makol et al. (2019) • Makol & Polo (2018) • Lerner et al. (2017) • Richters (1992)
3. Lack of models and strategies for testing the interpretability of divergence between informants' reports of adult personality and psychopathology	3. Examples of conceptual models and empirical strategies for directly testing the interpretability of divergence between informants' reports	• Al Ghriwati et al. (2018) • Cannon et al. (2020) • De Los Reyes et al. (2020) • De Los Reyes et al. (2013b) • De Los Reyes et al. (2009) • De Los Reyes & Ohannessian (2016) • Hartley et al. (2011)
4. Lack of consensus strategies for integrating comprehensive assessments of adult personality and psychopathology	4. Availability of measurement models for integrating multi-informant assessments and testing the validity of these models in relation to independent assessments	• Kraemer et al. (2003) • Makol, Youngstrom, et al. (2020)

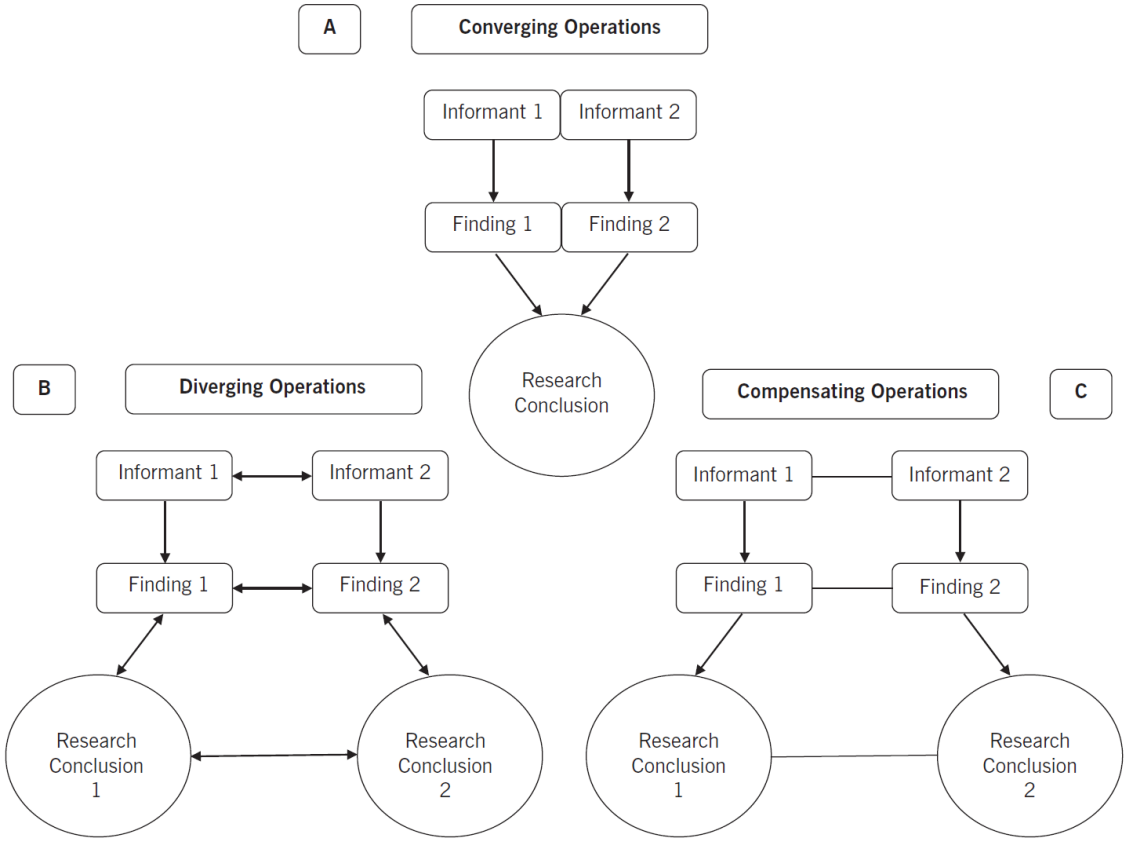

Figure 1. Graphical representation of the research concepts that comprise the Operations Triad Model. The top half (A) represents Converging Operations: a set of measurement conditions for interpreting patterns of findings based on the consistency within which findings yield similar conclusions. The bottom half denotes two circumstances within which researchers identify discrepancies across empirical findings derived from multiple informants' reports and thus discrepancies in the research conclusions drawn from these reports. On the left (B) is a graphical representation of Diverging Operations: a set of measurement conditions for interpreting patterns of inconsistent findings based on hypotheses about variations in the behavior(s) assessed. The solid lines linking informants' reports, empirical findings derived from these reports, and conclusions based on empirical findings denote the systematic relations among these three study components. Further, the presence of dual arrowheads in the figure representing Diverging Operations conveys the idea that one ties meaning to the discrepancies among empirical findings and research conclusions and thus how one interprets informants' reports to vary as a function of variation in the behaviors being assessed. Lastly, on the right (C) is a graphical representation of Compensating Operations: a set of measurement conditions for interpreting patterns of inconsistent findings based on methodological features of the study's measures or informants. The dashed lines denote the lack of systematic relations among informants' reports, empirical findings, and research conclusions. Originally published in De Los Reyes, Thomas, et al. (2013).

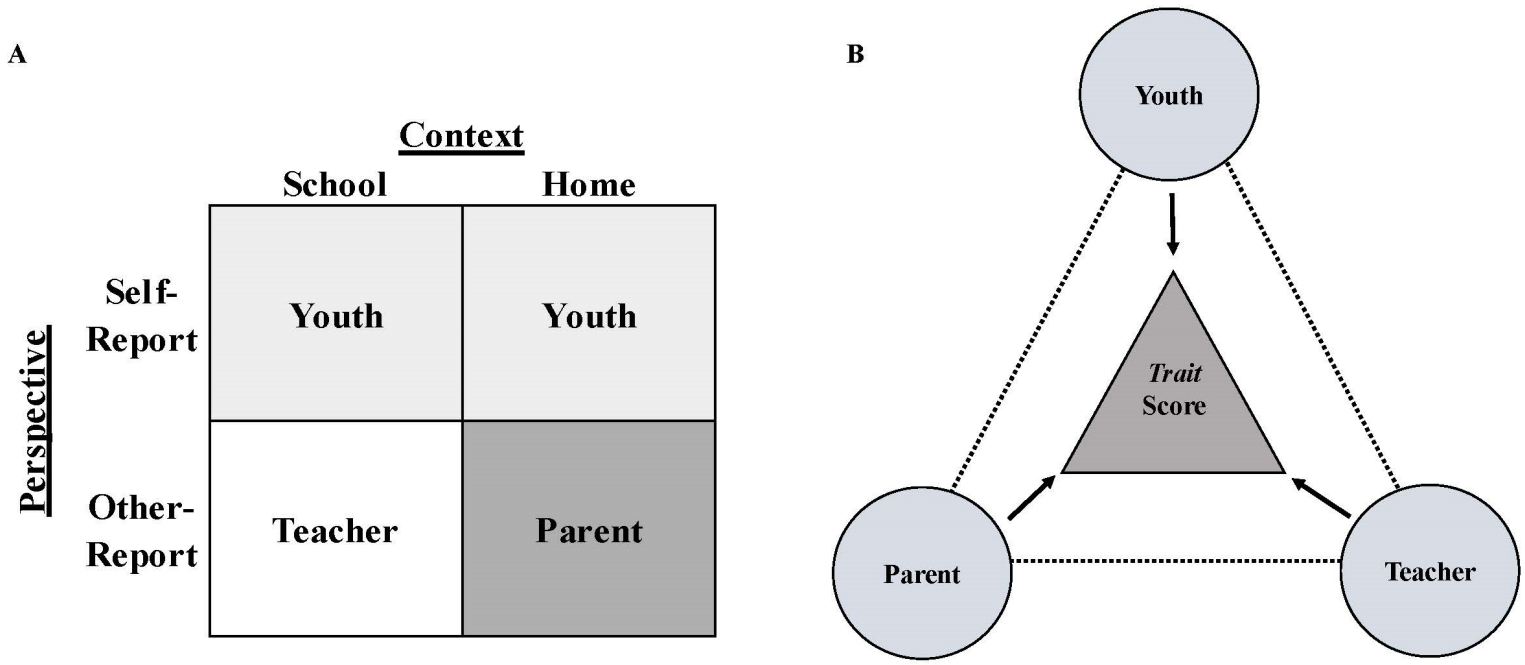

Figure 2. From Makol, Youngstrom, and colleagues (2020): **Panel A:** Example of use of "mix-and-match" criterion to identify optimal informants to include in a multi-informant assessment. Informants systematically vary in the perspective and context from which they rate youth mental health symptoms, with the goal of effectively triangulating on a *Trait* score. **Panel B:** Graphical depiction of multi-informant reports triangulating, much like GPS, to identify the *Trait* score. Both teacher- and parent-report provide information from an other-perspective, with teachers providing information about the school context and parents providing information about the home context. Youth reports provide the self-perspective and information about both the school and home contexts. Figures adapted from Kraemer et al. (2003).

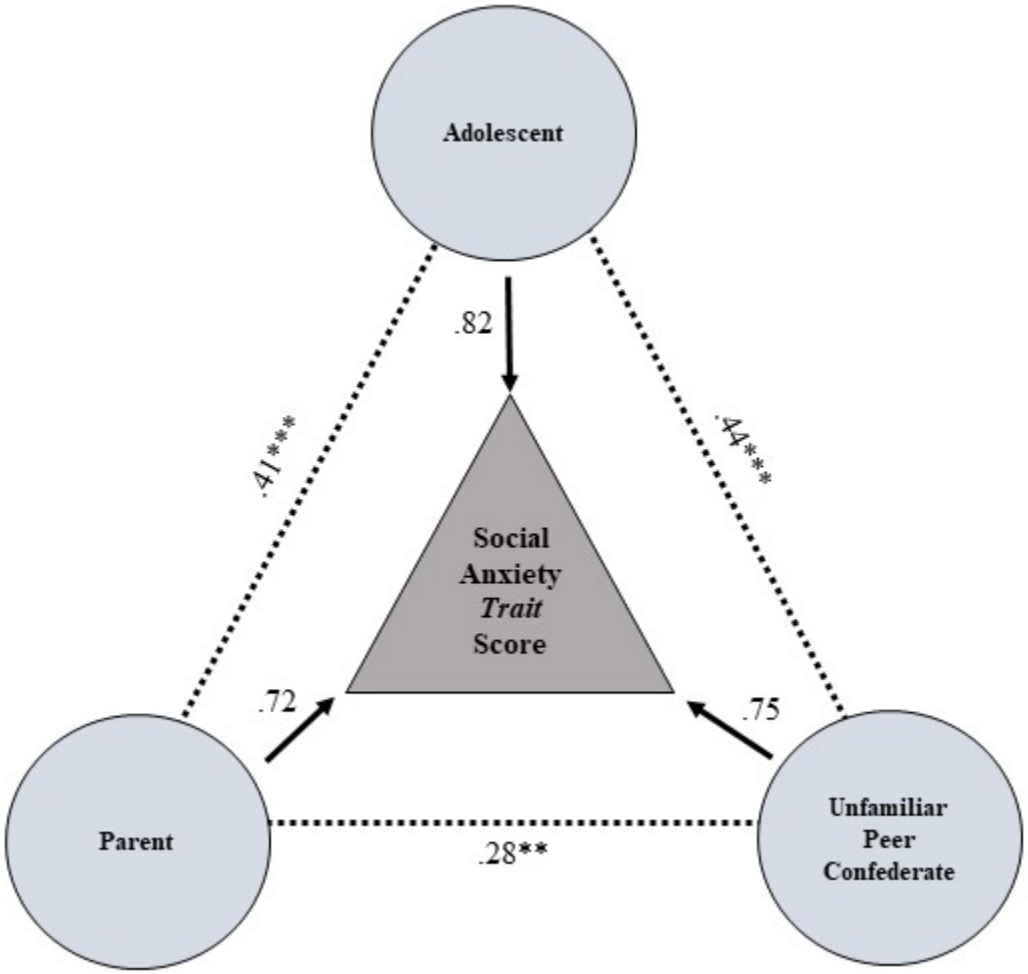

Figure 3. From Makol, Youngstrom, and colleagues (2020): Values on dotted lines denote bivariate correlations among parent, adolescent, and unfamiliar peer confederate reports of adolescent social anxiety. Values on solid arrows denote component weights from principal components analysis of parent, adolescent, and unfamiliar peer confederate reports. $^{**}p < .01$; $^{***}p < .001$.