

**EDITORIAL**

**Introduction to the Special Section:  
The Research Domain Criteria's Units of Analysis and  
Cross-Unit Correspondence in Youth Mental Health Research**

Andres De Los Reyes<sup>1</sup>

Deborah A. G. Drabick<sup>2</sup>

Bridget A. Makol<sup>1</sup>

Rafaella J. Jakubovic<sup>2</sup>

<sup>1</sup>University of Maryland at College Park

<sup>2</sup>Temple University

**CITATION**

De Los Reyes, A., Drabick, D.A.G., Makol, B.A., & Jakubovic, R. (2020, May).

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

### **Author Notes**

Andres De Los Reyes and Bridget A. Makol, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland, College Park, MD.  
Deborah A. G. Drabick and Rafaella J. Jakubovic, Department of Psychology, Temple University, Philadelphia, PA.

The first and third authors' work on this paper was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A180032 to University of Maryland at College Park. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Correspondence regarding this manuscript should be addressed to Andres De Los Reyes, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland, Biology/Psychology Building, Room 3123H, College Park, MD 20742; Office: 301-405-7049; Fax: 301-314-9566; E-mail: [adlr@umd.edu](mailto:adlr@umd.edu); Twitter: [@JCCAP\\_Editor](https://twitter.com/JCCAP_Editor)

### Abstract

In 2010, the National Institute of Mental Health launched the Research Domain Criteria (RDoC). The RDoC seeks to enhance research on the “active ingredients” of mental health concerns, and conceptualizes these concerns as disorders of neural circuitry. A key focus of the RDoC involves understanding mental health across biopsychosocial domains that traverse traditional diagnostic categories, such as impairments in affect, cognition, regulatory systems, and social processing. The RDoC advanced a framework for understanding these impairments as they manifest across multiple *units of analysis*: measured variables for which activity reflects functioning across biopsychosocial systems (e.g., subjective reports, behavioral performance, neural activity, cellular assays). However, scholars whose work focuses on children and adolescents encounter challenges with adopting the RDoC’s units of analysis framework, which requires a developmental adaptation. In particular, 50 years of research points to a key reality: the RDoC’s units of analysis will yield low correspondence, and this low correspondence will translate to inconsistent research findings. Yet, no guidelines exist for interpreting low correspondence among these units or integrating them for research purposes. Prior work on clinical assessments of children and adolescents points to theoretical and methodological models that facilitate interpreting and integrating multi-source assessments. In this introductory article to an RDoC-informed Special Section, we discuss how researchers might apply these models to interpreting and integrating the RDoC’s units of analysis using a developmental perspective. This Special Section reveals new lines of RDoC-informed research, particularly the development of paradigms for *strategically* integrating activity from multiple units of analysis.

**Keywords:** correspondence; Operations Triad Model; Research Domain Criteria; units of analysis

Mental health concerns incur considerable personal, societal, and financial costs and affect tens of millions of people worldwide (e.g., Kazdin & Blasé, 2011). To facilitate identifying, classifying, and treating mental health concerns, mental health professionals (we use “mental health professionals” throughout when we wish to refer to practitioners and researchers altogether) developed nosological systems to promote the reliable detection of these concerns and facilitate treatment planning (e.g., *Diagnostic and Statistical Manual of Mental Disorders [DSM-5]*; American Psychiatric Association [APA], 2013; *International Classification of Diseases [ICD-10]*; World Health Organization, 2007). Mental health professionals also developed a host of reliable and valid instruments to assess mental health concerns and the factors that may be associated with risk or resilience for these issues (Hunsley & Mash, 2007). Collectively, these instruments capably assess concerns across developmental periods, inform various aspects of clinical decision-making (e.g., screening, diagnosis, treatment planning, and treatment response assessments), and encompass a range of modalities (e.g., survey, interview, observational, and performance-based indices).

Despite these efforts and advancements, mental health professionals encounter considerable challenges when administering reliable and valid mental health assessments to children and adolescents (referred to as “youth” unless otherwise specified) and making sound decisions based on these assessments. Chief among these challenges are (a) current assessment technologies require extensive training and supervision to administer, score, and interpret (Mash & Hunsley, 2005); (b) clients often experience concerns that result in multiple diagnoses (Drabick & Kendall, 2010); (c) clients may display concerns in one or more social contexts (e.g., home, school, peer interactions; De Los Reyes et al., 2015); and (d) no one “gold standard” index exists for assessing and classifying any one domain of mental health concerns (De Los Reyes,

Augenstein, & Aldao, 2017; Richters, 1992). These challenges create uncertainties in clinical decision-making, thus resulting in impediments for both clinical practice (e.g., treatment planning and monitoring treatment response) and basic research (e.g., identifying mechanisms underlying mental disorders).

Ten years ago, the National Institutes of Health launched the Research Domain Criteria (RDoC; Insel et al., 2010). The RDoC is a research initiative designed to identify aspects of functioning that cut across current diagnostic categories (e.g., anxiety and mood disorders) and promote translational research that increases the efficacy of mental health treatment and prevention programs. Specifically, the RDoC initiative conceptualizes mental disorder symptoms and diagnoses as dysfunctions of brain circuitry that result in observable and thus measurable impairments. Within the RDoC initiative, these impairments fall into five functional domains that cut across traditional diagnostic categories (e.g., anxiety and mood, autism spectrum, psychoses). These domains include positive affect, negative affect, cognition, social processing, and regulatory systems (Sanislow et al., 2010). Of particular interest to mental health professionals is how the RDoC initiative seeks to discover mechanisms underlying dysfunctions of brain circuitry. Indeed, one of the RDoC initiative's key premises is that if we understand the mechanisms underlying mental disorders, we can develop focused treatments or refine existing treatments to specifically target these mechanisms. If successful, the RDoC initiative will result in more effective mental health interventions than those currently available. In this way, the RDoC initiative will potentially contribute to considerable improvements in public health.

Studies guided by the RDoC initiative seek to understand dysfunction(s) within and across the RDoC domains using multiple outcome measures. These multiple outcome measures reflect activity along multiple *units of analysis*. Originally developed in reference to mental

health research with adults, the RDoC's units of analysis include self-reports, behavior, physiology, neural circuitry, cells, molecules, and genes (Cuthbert & Insel, 2010). The comprehensive nature of these units of analysis signals that, in theory, RDoC-informed studies should increase our understanding of the *whole person*. That is, these studies should include comprehensive assessments of functional activity within or across domains and methods, ranging from the most sensitive of biological assays to behavioral observations, performance-based tasks, and well-established clinical scales traditionally administered in clinical research and practice settings. Presumably, these units of analysis vary both in the means by which they assess activity (i.e., direct observation of symptoms vs. assays of biological functioning) and the degree to which the activity assessed reflects functioning in one or more biopsychosocial systems.

The RDoC initiative has profoundly impacted mental health research in the past decade. In fact, a recent Google Scholar search of the seminal paper initially describing the RDoC initiative (Insel et al., 2010) reveals that this piece has received nearly 4,000 citations (search conducted February 1, 2020). Yet, three barriers exist to effectively applying the RDoC initiative to research with youth.

First, the RDoC initiative's units of analysis do not account for long-held traditions in youth mental health assessments. Consider that the focus on multiple units of analysis represents a significant innovation in research on adult mental health. Clinical assessments of adults have historically relied on clients' self-reports (for a review, see Achenbach, Krukowski, Dumenci, & Ivanova, 2005). However, multi-unit approaches to measurement have long-comprised core components of best practices in evidence-based assessment of youth mental health (Achenbach, McConaughy, & Howell, 1987; De Los Reyes, 2011; Kraemer et al., 2003). As such, key findings from studies of multi-unit assessments in youth mental health illustrate potential issues

when applying the RDoC initiative to research with youth. For example, clinical assessments of youth involve collecting clients' self-reports as well as reports from additional *informants*, typically significant others in clients' lives such as parents and teachers (Hunsley & Mash, 2007). Over 50 years of research and several hundred studies indicate relatively low levels of correspondence among ratings taken from these multiple informants' subjective reports (e.g., De Los Reyes et al., 2015). In fact, these correspondence levels are so low that no one informant's subjective report is interchangeable with another informant's subjective report (Achenbach, 2006). Thus, with the "self-report" unit of analysis alone, the RDoC requires substantial modification to "fit" RDoC-informed research with youth. Similarly, the "behavior" unit of analysis insufficiently reflects the myriad approaches to reliably and validly assess youth using behavioral measures, which include controlled observations, official records, and results from performance-based tasks. Consequently, a key aim of this Special Section was to advance a modified framework for understanding units of analysis relevant to RDoC-informed research with youth. This new framework for units of analysis will subsequently inform methods for addressing the additional aims described below.

Second, within the RDoC's framework, mental health professionals should understand functioning along *multiple* units of analysis. To understand how the RDoC initiative might ultimately inform clinical research and practice with youth, we must examine whether RDoC-informed studies leverage multiple units of analysis, and if so, which ones they use. In this Special Section, we apply the developmentally modified units of analysis framework described in this article to quantitatively examine RDoC-informed research with youth to date. Specifically, contributors to this Special Section applied the framework to quantitatively examine frequencies of units of analysis used to address aims in RDoC-informed research.

Third, we mentioned previously that a robust observation in youth mental health research is that reports from multiple informants tend to yield relatively low levels of correspondence (Achenbach et al., 1987; De Los Reyes et al., 2015). These low levels of correspondence often arise because (a) youth vary in the specific contexts in which they display signs of mental health concerns and (b) multiple informants (i.e., units of analysis) often vary in the contexts in which they have the opportunity to observe clients (e.g., parents at home vs. teachers at school; for a review, see De Los Reyes, Thomas, Goodman, & Kundery, 2013). However, this focus on multi-informant assessments means that our understanding of multi-source correspondence is largely restricted to activity relevant to the unit of analysis dealing with *subjective* reports, currently referred to as the “self-report” unit of analysis. Our Special Section provides a unique opportunity to not only report meta-analytic data on correspondence among the RDoC’s units of analysis, but also to identify factors that moderate levels of such correspondence. Indeed, in prior work with multiple informants’ subjective reports, key moderators of cross-unit correspondence include the observability of the domain of functioning assessed (e.g., greater correspondence for ratings of aggressive behavior vs. anxiety and mood concerns) and the overlap in contexts of observation between informants (e.g., greater correspondence for ratings completed by pairs of parents vs. a parent and a teacher; Achenbach et al., 1987; De Los Reyes et al., 2015). Thus, a key principle of correspondence among multiple subjective reports is that the observability of a domain of mental health functioning may affect the level of correspondence between units of analysis assessing that domain. Does this same principle generalize to the RDoC initiative’s units of analysis? Addressing this question in relation to the RDoC initiative will provide researchers and practitioners with an evidence base for interpreting correspondence along the RDoC’s units of analysis as indicators of mental health across developmental periods. In turn, we expect the

availability of this evidence base to reduce uncertainty in both decision-making within research and practice settings, and planning future research that addresses gaps identified by the meta-analysis included in this Special Section.

### **Purpose of the Special Section**

The purpose of this introductory article to the Special Section is four-fold. First, we advance a framework for understanding the RDoC's units of analysis as employed in research among youth. Second, we provide an overview of research on multi-source correspondence relevant to the RDoC's units of analysis. Third, we discuss theoretical and methodological models—developed and tested largely within research on informants' subjective reports of youth mental health—that may facilitate interpreting and integrating the RDoC's units of analysis. Fourth, we provide an overview of the articles in this Special Section, with particular attention paid to a meta-analysis that illustrates use of the developmentally modified framework for the RDoC's units of analysis to understanding a specific RDoC domain (i.e., social processing; Clarkson et al., 2020). We expect articles in this Special Section to not only inform strategies for selecting and integrating units of analysis in RDoC-informed research, but also serve as an important resource for hypothesis generation.

### **Developmentally Modified Framework for Understanding the RDoC's Units of Analysis**

In preparing this Special Section, we took a developmentally informed approach to the RDoC's units of analysis. As mentioned previously, a key goal of the Special Section involved taking a meta-analytic approach to estimating levels of correspondence among units of analysis relevant to RDoC-informed research with youth (Clarkson et al., 2020). Best practices in quantitative reviews that are focused on estimating levels of correspondence among disparate measures involves coding the sources that completed the measures (e.g., informant, measurement

modality; see Achenbach et al., 1987; De Los Reyes et al., 2015; De Los Reyes, Lerner et al., 2019; Duhig, Renk, Epstein, & Phares, 2000; Stratis & Lecavalier, 2015). In this way, quantitative reviews of multi-unit correspondence can test whether the measurement source explains variability in levels of correspondence among measures. Similarly, levels of correspondence among units of analysis may change as a function of the source of unit measurement. We previously discussed the mismatch between the current RDoC units of analysis framework and units of analysis as administered in youth mental health assessments. This required developmentally adapting the RDoC's units of analysis framework to "fit" RDoC-informed research with youth. We present this adapted framework in Table 1. Our adaptations followed three developmentally informed principles. We describe each of these principles below, both in how they informed adapting the framework, and their implications for interpreting the RDoC's units of analysis.

### **Developmental Principle #1: Subjectivity Rests in the Eyes of Multiple Beholders**

As of this writing, if one accesses the RDoC's main webpage regarding units of analysis (National Institute of Mental Health, 2020), the description of the unit of analysis focused on subjective impressions of domain-relevant functioning reads as follows: "Self-Reports refer to interview-based scales, self-report questionnaires, or other instruments that may encompass normal-range and/or abnormal aspects of the dimension of interest." This definition highlights many *methods* for gathering subjective impressions of domain-relevant activity. Yet, the definition ascribes little attention to the *informant* from whom one measures subjective impressions.

Within the subjective unit of analysis, the lack of attention to the information source may stem from the RDoC's focus on research with adults and concomitant reliance on self-reports.

Yet, here too the operational definition would be surprisingly limited. Indeed, over the last 15 years, assessments of adult mental health have increasingly leveraged reports from not only clients and trained assessors (e.g., clinical interviewers), but also significant others in clients' lives such as spouses, and in the case of aging adults, their caregivers (e.g., Achenbach et al., 2005; Achenbach, 2006; Hunsley & Mash, 2007; Vazire, 2006). As mentioned previously, for decades clinical assessments of youth have incorporated subjective data from multiple informants (De Los Reyes, 2011). Further, the grand majority of data supporting evidence-based intervention services for youth stems from multi-informant assessments of intervention outcomes (e.g., De Los Reyes et al., 2017; Weisz, Jensen Doss, & Hawley, 2005), which consequently justifies developmentally adapting the RDoC's units of analysis framework for research with youth. However, the need for developmentally adapting the RDoC's focus on subjective units of analysis goes well beyond simply documenting the fact that clinical assessments that include subjective reports are not limited to clients' self-reports. Indeed, as mentioned previously, decades of research indicate that subjective measures taken from different informants are not interchangeable (Achenbach et al., 1987; De Los Reyes et al., 2015). Further, expanding upon subjective units of analysis considered within RDoC-informed research allows for consideration of an emerging body of work focused on innovative models for interpreting and integrating multi-informant and multi-modal data (e.g., De Los Reyes, Cook, Gresham, Makol, & Wang, 2019; Drabick & Kendall, 2010; Kraemer et al., 2003). As detailed below, these models may inform strategies for interpreting and integrating units of analysis in RDoC-informed research among youth.

**Developmental Principle #2: Behavioral Indices Vary Considerably in Form and Referent**

Our second developmental principle holds that the unit of analysis focused on indexing behavior requires special attention paid to the form of the behavioral index, as well as the degree to which domain-relevant activity from an index reflects specific environmental influences on behavior. Decades of basic research on youth mental health supports a key idea: contextual factors influence youth in reciprocal and transactional ways, and can thus serve a protective function, as well as confer risk or exacerbate psychosocial difficulties (e.g., Drabick & Kendall, 2010; Goodman & Gotlib, 1999; Granic & Patterson, 2006). Further, environmental contingencies that exacerbate or minimize displays of domain-relevant behaviors may change considerably, depending on contextual demands, expectations, and influences (Skinner, 1953). Thus, a core feature of many evidence-based interventions for youth involves either manipulation of, or fostering adaptations to, environmental risk and protective factors for mental health concerns. Indeed, established psychosocial interventions focus on changing or fostering adaptations to such environmental factors as parenting (e.g., parent management training), peer relations (e.g., social skills training), anxiety-provoking stimuli (e.g., exposure-based therapies), and interpersonal relationships (e.g., interpersonal psychotherapy; for reviews, see Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; Weisz, Hawley, & Jensen Doss, 2004).

It logically follows that a behavioral measure for which environmental factors proximally impact scores may produce clinically meaningful differences in domain-relevant activity levels, relative to a measure for which scores have distal referents to environmental factors. For example, consider an assessment that incorporates two behavioral tasks designed to measure processes relevant to disruptive behavior in young children. Perhaps one task focuses on assessing processes that manifest within the context of the parent-child relationship, such as a

parent-child interaction task designed to measure parental warmth. That task may produce meaningfully distinct results relative to the results from a task designed to measure processes distally impacted by environmental factors present earlier in the child's development, such as an inhibitory control task for which a child experiences performance deficits stemming from prenatal exposure to substances. Further, between any two behavioral measures expected to be impacted by environmental factors, whether the environmental referents vary might also produce clinically meaningful differences in estimates of domain-relevant activity (e.g., standardized observations of children in home vs. school contexts or among adolescents engaged in problem-solving discussions with peers vs. parents; Dishion, Nelson, & Bullock, 2004). Thus, RDoC-informed research with youth ought to pay special attention to whether variability in activity levels indexed by behavioral measures reflects proximal responses to or distal influences by environmental factors.

**Developmental Principle #3: Biological Indices Vary Not Just in Their Reflections of Disparate Systems, but Whether They Index Context-Dependent Processes**

The first two developmental principles point to the idea that variations in how one assesses domain-relevant activity within the subjective and behavioral units of analysis may produce substantial variations among estimates of domain-relevant activity within these "classes" of units. Such is likely the case with biologically based units of analysis within the RDoC initiative: physiology, circuits, cells, molecules, and genes. Across these units, one finds examples of measures that require a stimulus or social experience to elicit activity levels on that measure. For example, female adolescents assessed using the Taylor Aggression Paradigm, which involves competition with and setting a level of electric shock for a fictitious opponent, evidence increased autonomic arousal and effects of cortisol on heart rate in the context of an

aggression induction relative to a non-aggression control context (Rinnewitz et al., 2019). As a second example, relative to healthy controls, youth with social anxiety disorder exhibit greater fear conditioning (i.e., indexed by the startle-blink reflex) in the context of negative unconditioned stimuli (i.e., negative insults and critical faces) compared to neutral and positive stimuli (Lissek et al., 2008). Additionally, research on biology  $\times$  environment interactions suggest that both factors are important to consider in understanding youth functioning, even if the contextual influences are distal to the phenomena of interest (Beauchaine et al., 2008). For instance, relative to youth insecure attachments, youth with secure attachments and long serotonin transporter alleles evidence less stress, as indexed by electrodermal activity, during the Trier Social Stress Test for Children, which involves giving a speech and completing mental arithmetic tasks in front of an audience (Gillissen et al., 2008). Similarly, prior work discovered an interaction between peripheral serotonin levels and mother-adolescent interactions in predicting self-harm behavior (Crowell et al., 2008). In this study, girls with relatively high peripheral serotonin were at risk for self-harm in the context of highly negative mother-adolescent interactions, whereas girls with relatively low peripheral serotonin were at risk for self-harm regardless of the mother-adolescent interaction context.

In contrast, other biological measures estimate domain-relevant activity using procedures for which the context in which one takes the measure has little-to-no impact on the results obtained. For instance, performance on tasks that assess aspects of executive functioning (e.g., working memory, set-shifting), which are thought to reflect functioning of the dorsolateral prefrontal cortex among other areas (Vanderhasselt et al., 2006, 2007), should not be affected by contextual influences beyond those associated with task procedures. Nevertheless, although proximal contextual factors may not affect performance on some biological indices, distal factors

modify biological indices in ways that would affect youth's performance even at baseline. Evidence from genotype-environment correlations indicates that youth may be at risk for psychosocial difficulties because caregivers typically provide factors that contribute to *both* genetic and contextual risk. For example, youth who are reared in high-risk environments (e.g., characterized by physical discipline, family conflict, or community violence) may exhibit resilience from psychological symptoms because of the presence of functional polymorphisms (e.g., with MAOA; Edwards et al., 2010) or their capacity for emotion regulation, which can be indexed by autonomic reactivity (Beauchaine, 2008; Bubier, Drabick, & Breiner, 2009). Needless to say, methodological distinctions among these groups of measures, such as the reliability of the instrument, may explain some of the differences in results obtained. Yet, a substantial amount of variability in estimates obtained from different biological measures may reflect the degree to which the biological process indexed by each measure changes across contexts (i.e., reflecting proximal contextual influences), or alternatively displays little change across contexts (i.e., reflecting distal contextual influences). In light of both basic and applied research indicating the substantial influence that environmental factors have on the development, attenuation, exacerbation, and maintenance of youth mental health concerns, RDoC-informed research with youth must carefully attend to the context-sensitivity of units of analysis reflecting biological processes.

### **The Ubiquity of Low Correspondence among Units of Analysis**

We briefly reviewed a set of developmentally informed principles that guided our adaptations to the RDoC's units of analysis framework. Collectively, the principles highlight two elements of the RDoC's units of analysis that have received relatively little attention. First, the units of analysis each harbor a wide diversity of indices that vary considerably in format as well

as the processes they were constructed to reflect. Second, the large variety of indices available dictates that no one index within a unit of analysis will produce estimates of domain-relevant activity that are interchangeable with estimates from another index within that same “class” of unit of analysis (e.g., pair of behavioral indices, pair of circuit indices). The considerable degree of *within-unit variability* consequently leads us to expect considerable *between-unit variability* as well, a related issue that has also received relatively little attention in RDoC-informed research.

### **Correspondence among Multiple Informants’ Subjective Reports**

To what degree does the extant literature support the notion that scholars conducting RDoC-informed research should expect low correspondence in estimates of domain-relevant activity between “classes” of units of analysis? The strongest evidence exists in studies estimating correspondence levels between subjective reports completed by multiple informants, namely, parents, teachers, and youth themselves (Achenbach, 2006). We trace the first of these investigations back to the late 1950s (Lapouse & Monk, 1958). Nearly 40 years later, Achenbach and colleagues (1987) published a seminal meta-analysis of 119 studies indicating that the mean level of cross-informant correspondence fell in the low-to-moderate range ( $r = .28$ ). Our team observed that same effect ( $r = .28$ ) nearly 30 years later (De Los Reyes et al., 2015), in a sample of 341 studies published after Achenbach and colleagues (1987). One might argue that the stability of these effects rests, in part, on studies testing correspondence based on estimates taken from similar samples. For example, perhaps all the cross-informant correspondence studies originate from North America, using a restricted set of standardized instruments completed within a homogeneous cultural context (see also Henrich, Heine, & Norenzayan, 2010). This is not so; indeed, in a recent meta-analysis we coded for country of origin in a sample of 314 studies and found stable, low-to-moderate levels of cross-informant correspondence across 30

countries and six continents, thus illustrating the global significance of cross-informant correspondence (De Los Reyes, Lerner et al., 2019). Findings from cross-informant correspondence research are some of the most robust observed in the social sciences (Achenbach, 2006); they rival those observed for placebo effects (cf. Ashar, Chang, & Wager, 2017). Thus, if levels of correspondence remain robustly low-to-moderate for data collected across different “types” of subjective reports, this likely signals that we should expect even lower levels of correspondence between different units of analysis.

### **Correspondence between Subjective Reports and Biobehavioral Units of Analysis**

Issues of cross-unit correspondence have received relatively less attention than the multi-informant correspondence work reviewed previously. Yet, patterns of correspondence within and across biobehavioral units of analysis likely occur at even lower magnitudes than estimates within subjective reports of youth mental health. On average, multi-informant (e.g., youth, parents, teachers) reports of youth behavior are associated at only small-to-moderate levels with observed behavior, including for externalizing problems (e.g.,  $r_s = .01-.52$ ; Becker, Luebke, Fite, Stoppelbein, & Greening, 2014; Doctoroff & Arnold, 2004; Henry, 2006; Hinshaw, Han, Erhardt, & Huber, 1992; Hymel, Rubin, Rowden, & LeMare, 1990; Wakschlag et al., 2008; Winsler & Wallace, 2002), internalizing problems (e.g.,  $r_s = .01-.48$ ; Becker et al., 2014; Beidel, Turner, Hamlin, & Morris, 2000; Cartwright-Hatton, Hodges, & Porter, 2003; Hinshaw et al., 1992; Glenn et al., 2019; Hymel et al., 1990; van Doorn, Bodden, Jansen, Rapee & Granic, 2018; Winsler & Wallace, 2002), and social functioning (e.g.,  $r_s = .02-.42$ ; Beidel et al., 2000; Cartwright-Hatton et al., 2003; Doctoroff & Arnold, 2004; Glenn et al., 2019; Hymel et al., 1990; Winsler & Wallace, 2002). Overall, the vast majority of these studies found less than small-magnitude correlations between informants’ reports and observed behavior ( $r_s < .20$ ), with

larger correlations observed when informants rated youth behavior occurring in the same context in which they typically observe youth (e.g., teacher reports of disruptive behavior and classroom observations of disruptive behavior; Doctoroff & Arnold, 2004; Winsler & Wallace, 2002).

Subjective reports also exhibit low convergence with physiological indices. For example, inconsistencies are found when comparing adolescent self-reports to biological measures of substance use (e.g., urine and hair biospecimens). Specifically, discrepant results between these two modalities occur within 11% to 42% of cases, with some findings indicating positive tests based on self-report and not biospecimens, and vice versa (Akinici, Tarter, & Kirisci, 2001; Delaney-Black et al., 2010; Dembo et al., 2015; Harris, Griffin, McCaffrey, & Morral, 2008; Lennox, Dennis, Ives, & White, 2006; Williams & Nowatzki, 2005). Larger discrepancies are observed between adolescent self-reports and biospecimens for more severe and less commonly used substances (e.g., cocaine, crack, opiates, hallucinogen, amphetamines), relative to more commonly used substances (e.g., alcohol, marijuana; Delaney-Black et al., 2010; Harris et al., 2008). These low levels of convergence between subjective reports and biospecimens have led to concerns about the validity of adolescent substance use assessments, and have been attributed to both intentionally inaccurate reports (e.g., due to social desirability bias or social stigma) and unintentionally inaccurate reports (e.g., due to poor recall; Harris et al., 2008).

As another example, consider anxiety assessments. Researchers studying youth anxiety commonly incorporate measures of physiological arousal (e.g., heart rate [HR], HR variability, skin conductance; De Los Reyes & Aldao, 2015). Youth with anxiety concerns perceive greater physiological arousal during anxiety-provoking tasks compared to youth without anxiety concerns, even when no group differences are observed in objective physiological measurements (Anderson & Hope, 2009; Miers, Blote, Sumter, Kallen, & Westenberg, 2011; Schmitz, Blechert,

Kramer, Asbrand, & Tushen-Caffier, 2012). Further, youth and parent reports of youth anxiety and physiological arousal are often uncorrelated or inconsistent with objective physiological measurements (Anderson, Veed, Inderbitzen-Nolan, & Hansen, 2010; Blom, Olsson, Serlachius, Ericson, & Ingvar, 2010; De Los Reyes et al., 2012; Miers et al., 2011; Greaves-Lord et al., 2010; Schmitz, Tushen-Caffier, Wilhelm, & Blechert, 2013), whereas other studies find low-magnitude correlations between subjective reports and objective physiological measurements (Greaves-Lord et al., 2010; Miers et al., 2011). What complicates these findings is the high variability found among objective physiological measurements, with factors such as type of physiological measurement (e.g., HR, HR variability), clinical status (i.e., low vs. high anxiety), and youth sex contributing to unique associations between subjective reports and objective physiological measurements. These findings have led to questions about how to best understand physiological processes among youth with anxiety, including how to best identify and characterize anxious youth from their non-anxious counterparts (e.g., using subjective or objective measurements, integrating subjective and objective information sources).

### **Correspondence among Biobehavioral Units of Analysis**

Relative to the deep literature on correspondence between subjective reports, and the relatively nascent literature on correspondence between subjective reports and other units of analysis (e.g., behavior and physiology), even less research considers correspondence among biobehavioral units of analysis (e.g., behavioral, physiological, neurological data). The lack of research attention on correspondence among these units may stem from the limited guidance available on interpreting convergence or divergence among these units (De Los Reyes & Aldao, 2015). In psychological research among youth in particular, attention must be paid to how the processes measured change over time, and the ability of our assessment measures to capture

constructs of interest (e.g., emotion regulation) across development (Adrian, Zeman, & Veits, 2011). Although there is generally agreement that youth mental health assessments should leverage multi-method approaches and incorporate multiple levels of analysis (Buss, Davidson, Kalin, & Goldsmith, 2004; Davis & Ollendick, 2006; Deveney et al., 2019; Drabick & Kendall, 2010; Leventon, Stevens, & Bauer, 2014), findings for correspondence across methodologies are mixed.

Consider that non-subjective report measures may index distinct constructs or levels of these constructs. For example, in youth emotion research, observational measures may accurately capture overt displays of emotion, whereas physiological measures may accurately capture covert or internal emotional experiences. Hubbard and colleagues (2004) collected observational and physiological data to index anger among children during a procedure in which they lost a game to a confederate who cheated. Observational variables included angry facial expressions and nonverbal behavior, and physiological variables included skin conductance reactivity and HR reactivity. Although the authors observed significant associations between skin conductance reactivity and HR reactivity, these two units differed in their relations to reactive aggression. That is, the study revealed relations between skin conductance reactivity and reactive aggression, but not between HR reactivity and reactive aggression. These discrepant findings may reflect the idea that skin conductance reactivity is a more direct reflection of emotional arousal, whereas HR reactivity may reflect additional attentional processes. Interestingly, although skin conductance reactivity was positively associated with angry facial expressions and nonverbal behavior, observational and physiological measures were not robustly related, with an association between angry nonverbal behavior and skin conductance reactivity only emerging among youth rated as displaying high levels of reactive aggression by their teachers.

As another example, consider that researchers examining fearful temperament among children using multiple biobehavioral units of analysis also report contradictory results. Buss and colleagues (2004) compared multiple measures of fearful temperament behaviors with measures of physiological reactivity, such as cortisol and cardiac reactivity, including HR, respiratory sinus arrhythmia (RSA), and pre-ejection period (PEP), among young children exposed to different mildly threatening contextual circumstances. Although investigators confirmed their hypothesis that freezing behavior during the stranger free-play scenario was associated with higher basal cortisol and resting PEP, they failed to replicate past findings that other behavioral indicators of fear (e.g., inhibition, facial fear, crying, escape behavior) are associated with higher cortisol and cardiac activity. Additionally, the investigators found no association between changes in HR, RSA, or PEP with any of the fear-related behavioral composite variables (all  $r$ s < .20).

Other research considering behavioral and physiological units of analysis reports discrepancies among these indicators as well. Miller et al. (2006) compared observed behavior, (i.e., classroom emotional expression and regulation), psychophysiological stress reactivity (i.e., HR reactivity and change in cortisol), and psychophysiological regulation, indexed by RSA, among low-income young children and reported modest associations between HR and displays of emotion in the classroom, and no associations between RSA and classroom behaviors. A study by Michels and colleagues (2013) further found moderate associations between HR variability and cortisol among children. Nevertheless, HR variability and cortisol index two different neural stress systems (the autonomic nervous system and hypothalamus-pituitary-adrenal axis, respectively) and differ in their reaction to stress, such that the HR variability reaction remains

high after several repeated stressors, whereas the cortisol stress response decreases through habituation.

Leventon, Stevens, and Bauer (2014) expanded upon studies comparing behavioral and physiological methods by including cross-unit comparisons with electrophysiological measures. These researchers assessed emotional responses via subjective ratings, event-related potentials (ERP), HR, and HR variability. Physiological measures may be used to measure emotional response over time and reflect the intensity of the response; for example, an initial deceleration in HR and increase in HR variability is evoked by moderately emotionally arousing stimuli (Leventon et al., 2014). ERP may also serve as a valuable assessment of emotion processing because of its sensitivity to real-time emotional arousal during an event (Leventon et al., 2014). The amplitude of the late-positive potential in ERP covaried with psychophysiological responses to negative, but not positive, stimuli. Further, this covariance was specific to the posterior cluster, and was stronger during the early window of the task ( $r = .40$ ). Taken together, these studies evidence divergence among biological and behavioral units of analysis when considering physiological and observational indices. However, there is a dearth of research considering some of the RDoC's units of analysis among children, such as genes with other units of analysis.

### **Theoretical and Methodological Models for**

#### **Interpreting and Integrating Multi-Informant Assessments**

Across several bodies of work, methodologically similar as well as distinct measures of youth mental health commonly yield estimates that, at best, display low-to-moderate correspondence with one another. In recent years, researchers studying correspondence levels among multiple informants' subjective reports have constructed models for directly testing whether levels of correspondence reflect clinically meaningful information. Conceptually

grounded empirical work provides compelling evidence that multiple informants' disagreements in reports of youth mental health signal individual differences in the specific contexts in which youth display mental health concerns. We briefly highlight two models, as they may inform paradigms for interpreting and integrating data derived from the RDoC's units of analysis.

### **Operations Triad Model**

Researchers conducting basic research on cross-informant correspondence have capitalized on key findings from the meta-analyses described previously. Specifically, the meta-analysis by Achenbach and colleagues (1987) found that when a pair of informants observe youth from the same context (e.g., two parents at home, two teachers at school), their reports tend to correspond to a greater extent than two informants who observe behavior from different contexts (e.g., parent, teacher). In this same meta-analysis, the correspondence levels tended to be higher for reports about externalizing (e.g., aggression, hyperactivity) than internalizing (e.g., anxiety, mood) concerns. These moderating factors were replicated in the more recent meta-analysis described previously (De Los Reyes et al., 2015).

Findings from meta-analyses reveal that the contexts in which informants observe behavior, and the observability of the behaviors about which they report, each contribute to disagreements between reports. Findings regarding disagreements between reports and the observability of the behaviors assessed converge with long-observed findings in research on factors affecting inter-rater reliability within assessments of observed behavior (e.g., Groth-Marnat & Wright, 2016). However, the idea that the contexts in which informants observe behavior might relate to disagreements between reports necessitated testing within controlled laboratory settings. The first laboratory-controlled test of these notions focused on examining the interpretability of parent and teacher reports of young children's disruptive behavior (De Los

Reyes, Henry, Tolan, & Wakschlag, 2009). The researchers leveraged a sample of 327 children whose parents and teachers provided disruptive behavior reports. The children and parents also participated in a laboratory observation task designed to assess the degree to which children displayed disruptive behavior within and across social contexts (Wakschlag et al., 2008).

Through a series of standardized tasks designed to elicit displays of relevant behaviors (e.g., tasks that include opportunities for the child to display noncompliance), independent observers coded for behaviors within and across interactions with parents (i.e., home context) and interactions with a non-parental clinical examiner (i.e., non-home contexts like school).

Consistent with the notion that variations in environmental contingencies (e.g., behavior of interaction partners) may produce alterations in behavior (e.g., Skinner, 1953), controlled laboratory observations revealed considerable variations in children's behavior. That is, children were distinguished by levels of disruptive behavior and based on the adult(s) with whom they exhibited this behavior: (a) high levels with both parent and examiner, (b) low levels with both parent and examiner, (c) high levels with only parent, or (d) high levels with only examiner. Further, patterns of parent and teacher reports were consistent with these variations in behavior within and across contexts. Specifically, those children for whom both parent's and teacher's subjective report endorsed disruptive behavior also tended to display disruptive behavior *across* parent and examiner interactions on the laboratory task. Additionally, those children for whom teacher's subjective report *but not* parent's subjective report endorsed disruptive behavior also tended to display disruptive behavior within examiner interactions *but not* parent interactions. The reverse was also true: those children for whom parent's subjective report *but not* teacher's report endorsed disruptive behavior tended to display disruptive behavior within parent interactions *but not* examiner interactions.

The controlled laboratory study by De Los Reyes and colleagues (2009) revealed that one could meaningfully interpret similarities and differences between parent and teacher reports of children's disruptive behavior as reflecting consistencies or inconsistencies in children's behavior across home and non-home contexts, respectively. Over the last decade, a considerable amount of research from multiple independent research teams has revealed similar findings in a variety of assessment settings (e.g., community, outpatient, inpatient, pediatric), mental health domains (e.g., internalizing, externalizing, neurodevelopmental), developmental periods of assessment (e.g., children, adolescents, adults), and informants (e.g., parents, teachers, self-reports, peers; for a review, see De Los Reyes, Cook et al., 2019). Several years ago, this body of evidence culminated in the development of a framework for conducting basic research on the interpretability of *patterns* among subjective reports. That is, do instances in which informants' subjective reports converge or diverge in estimates of behavior signal meaningful information about the behaviors being assessed? This is the key question addressed by research informed by the Operations Triad Model (De Los Reyes et al., 2013), which we depict in Figure 1.

The Operations Triad Model provides researchers with a guide for detecting various patterns of subjective reports. First, the model guides tests of whether instances in which subjective reports converge on the presence of a behavior signal consistencies in displays of the behavior across contexts (i.e., *Converging Operations*; Figure 1a). Second, when subjective reports diverge, the Operations Triad Model guides researchers through tests of whether these instances signal either meaningful inconsistencies in displays of the behavior across contexts (i.e., *Diverging Operations*; Figure 1b) or inconsistencies due to methodological confounds, such as differences among informants in terms of the instruments' psychometric properties (i.e., *Compensating Operations*; Figure 1c). To illustrate these principles, the study by De Los Reyes

and colleagues (2009) revealed evidence for *both* Converging Operations (i.e., agreement between reports signals cross-contextual displays of behavior) *and* Diverging Operations (i.e., disagreement between reports signals context-specific displays of behavior).

In sum, the Operations Triad Model facilitates building a basic science about those circumstances in which variations among subjective reports signal meaningful clinical information. Further, the Operations Triad Model has heuristic value in facilitating interpretations of measures beyond subjective reports of mental health. Specifically, modified versions of the model facilitate interpreting multi-informant assessments of family functioning (De Los Reyes & Ohannessian, 2016) and links between subjective reports and direct measures of physiology (De Los Reyes & Aldao, 2015). Consequently, we describe below how the Operations Triad Model might facilitate building a basic science of interpreting patterns of activity across the RDoC's units of analysis.

### **Kraemer's "Satellite Model"**

If the Operations Triad Model informs basic research pointing to the value of understanding variations among informants' subjective reports about behavior, how might one integrate multi-informant assessments to yield psychometrically sound estimates of behavior? This is the key question addressed by a model developed by Kraemer and colleagues (2003), which we graphically depict in Figure 2. The basis of the model lies in an innovative idea. Assessors often view the disagreements between informants' reports as a hindrance to interpreting data from these reports (see also De Los Reyes, 2011; De Los Reyes et al., 2015). What if assessors could turn these disagreements into a tool that *facilitates* how they interpret data from these reports? To address this question, Kraemer and colleagues (2003) turned to the geographic sciences, which commonly rely on multi-source data to make informed decisions

about the precise spatial location of objects and people. Detecting the spatial location of a *target*, like a building, involves collecting satellite data based, in part, on the latitude and longitude of the satellite's own location. One cannot arrive at precise location data for a target based on data from a single satellite, in part, because satellites rarely hover over the exact point of a target's location. For the same reasons, one cannot arrive at precise location estimates based on data from multiple satellites residing in the same or similar locations. Rather, precise location estimates of a target only result from leveraging data from multiple satellites that systematically vary in the latitude and longitude of their own locations, relative to the target. In this way, the satellites *triangulate* on the relative position of the target.

In the satellite metaphor described by Kraemer and colleagues, triangulation of multi-source data is the “active ingredient” in arriving at precise estimates of a target's location. For satellites, one achieves triangulation by “mixing and matching” satellites that systematically vary along two domains relevant to their utility in estimating a target's location, namely their latitude and longitude relative to the target. For example, one might achieve triangulation by relying on three satellites. Two of the satellites reside at similar latitudes but distinct longitudes. A third satellite resides in a latitude that differs from the two other satellites, but a longitude that “sits” in between the longitudes of the two other satellites. This strategic positioning of the satellites optimizes their use in triangulating on the location of the target.

Kraemer and colleagues (2003) innovatively applied this “satellite model” to use of multi-informant assessments. Based on meta-analytic research described previously, Kraemer and colleagues posited that the contexts in which informants observed the youth's behavior and the perspectives from which they perceived the youth could serve as the metaphorical equivalent of the latitude and longitude domains for satellites triangulating on a target (see Figure 2a). Thus,

parents and teachers perceive youth from an observer-perspective but from different contexts that require and thus elicit different behaviors (i.e., home vs. school). Conversely, the youth perceives their own behavior from a self-perspective, and bases their reports on observations of their behavior as it manifests across the contexts upon which parents and teachers base their reports. As depicted in Figure 2b, parents and teachers reside in the “same latitude” of perspective but “different longitudes” of contexts. Further, relative to parents and teachers, youth reside in a “different latitude” of perspective but a “longitude” of context that “sits” in between that of parents and teachers. By applying principal components analysis (PCA) to these reports, Kraemer and colleagues (2003) derived two component estimates for variability explained by *perspective* and *context*. A third *trait* component yielded an integrated behavioral score that takes into account the triangulated positioning of data across the three informants. Kraemer and colleagues (2003) surmised that a triangulated *Trait* score optimally estimates youth behavior.

In a recent test of the Kraemer and colleagues’ (2003) model, Makol and colleagues (2020) relied on three informants who provided reports of adolescent social anxiety. Assessments of social anxiety commonly rely on parent reports, given their involvement in soliciting services on behalf of their adolescent (Hunsley & Lee, 2014), and adolescent reports, given the value of understanding clients’ lived experiences with social anxiety (De Los Reyes & Makol, 2019). Importantly, individuals experiencing social anxiety concerns manifest them to a considerable degree when interacting with unfamiliar people (e.g., Cannon et al., 2020; Hofmann et al., 1999; Raggi et al., 2018). Thus, Makol and colleagues (2020) relied on a third informant, an *unfamiliar peer confederate*, who based their report on their observations of the adolescent within a series of controlled laboratory social interactions. All three of these informants provide psychometrically sound subjective reports about adolescent social anxiety, but do so in

systematically different ways (Deros et al., 2018; Glenn et al., 2019). As depicted in Figure 3, these three informants' reports triangulate in estimates of adolescent social anxiety, as indicated by the patterns of cross-informant correlations and component weights derived from a principal components analysis of the three reports constructed by Makol and colleagues (2020). Parents and peer confederates each perceive the adolescent from an observer perspective but from different contexts (i.e., home vs. non-home). Adolescents perceive their behavior from a self-perspective and across the home and non-home contexts covered by parents and peer confederates, respectively.

Makol and colleagues (2020) found that the *Trait* score estimate from the three informants' reports optimized prediction of clinically relevant criterion variables. Specifically, in this same sample, independent observers coded adolescent social anxiety as displayed in the social interactions between adolescents and peer confederates, and parents varied as to whether they sought a clinical evaluation on behalf of their adolescent (*clinic-referred*) or participated in a non-clinic evaluation of family relationships (*community control*). Importantly, these two criterion variables indexed behavior to a specific context (e.g., peer interactions) or information source (i.e., parent). Nevertheless, the *Trait* score estimate outperformed all three informants in their ability to predict the criterion variables and yielded large-magnitude predictions of observed anxiety ( $\beta$ s: .47-.67) and referral status ( $OR$ s: 2.66-6.53). As described previously, researchers typically observe rather small correlations between informants' reports about internalizing concerns and independent assessments of observed internalizing behaviors. Collectively, the Operations Triad Model (De Los Reyes et al., 2013) and the Kraemer and colleagues (2003) "satellite model" inform the development of paradigms for understanding and integrating multi-informant assessment data, and can be extended to integrate the RDoC's units of analysis.

**Applying Theoretical and Methodological Models of  
Multi-Informant Assessments to the RDoC's Units of Analysis**

Applying the Operations Triad Model and Kraemer's "satellite model" to the RDoC will require research seeking to identify the "latitudes and longitudes" of units of analysis. Stated another way, what factors might explain variability in domain-relevant activity among units of analysis? Two factors might be particularly useful to consider. When considering RDoC-informed research with youth, we would be remiss if we did not consider a key principle in developmental psychopathology, namely that mental health concerns arise out of the combined influence of biological, psychological, and sociocultural factors that make one susceptible to, or buffer against, developing maladaptive reactions to social contexts (Cicchetti, 1984). Yet, not all contexts affect youth the same way: If contexts may differ in the contingencies that give rise to behavior, then not all contexts elicit mental health concerns to the same degree and for all individuals (e.g., Carey, 1998; Kazdin & Kagan, 1994; Mischel & Shoda, 1995).

In this respect, one factor that might explain variability among the RDoC's units of analysis is the degree to which activity among units of analysis depends on stimuli that serve as referents to proximal influences of specific contexts. Examples might include tasks that elicit socio-evaluative threat (e.g., speech tasks; see Gunnar, Talge, & Herrera, 2009) or social exclusion (e.g., cyberball; Williams, Cheung, & Choi, 2000). For some units (e.g., indices of peripheral physiology or neural circuitry), activity levels change markedly in reference to specific stimuli; indeed, pulling the stimulus out of the model changes how one interprets activity levels. For other units, although social contexts might heavily influence activity levels, these influences are distal in nature. For instance, a blood test for detecting domain-relevant inflammatory or genetic markers may yield the same result regardless of the context in which

one takes the blood test. Thus, *stimulus dependence* may explain some of the variability in domain-relevant activity.

A second factor traces back to our discussion of moderators of cross-informant correspondence. Informants' subjective reports about mental health tend to yield greater levels of correspondence when the behavior being assessed contains more observable signs. For example, subjective reports about hyperactive behavior tend to correspond to a greater extent than subjective reports about worry (Achenbach et al., 1987; De Los Reyes et al., 2015). To be clear, we do not necessarily envision this exact internal-external pattern to manifest among the RDoC's units of analysis. In fact, high rates of comorbidity, even among current diagnostic categories, inspired the RDoC initiative's focus on functional domains that cut across traditional diagnostic boundaries. That being said, units for which activity levels depend on assessing processes with *external referents* (e.g., social skills deficits, behavioral avoidance) might differ in meaningful ways from units for which activity levels depend on assessing processes with *internal referents* that manifest "under the skin" (e.g., neural or HPA activity).

Taken together, we can surmise that another set of "latitudes and longitudes" explaining variability in the RDoC's units of analysis are that of *stimulus* and *process*, which we depict in Figure 4a. These two factors inform the strategic selection of units of analysis that triangulate on estimating domain-relevant activity, which we depict in Figure 4b. Consider the RDoC's negative valence domain. Several units of analysis capably assess activity in this domain. Applying the model of Kraemer and colleagues (2003) to this circumstance involves selecting units of analysis to *force* disagreement in activity levels among units that assess negative valence. For example, one might select two units that involve assessing activity that depends on a stimulus, such as threat. These two units might vary in the process they assess within threat-

relevant contexts, such as neural activity (internal) when exposed to a threat-eliciting task during neuroimaging, and social skills (external) when exposed to an anxiety-provoking social interaction task (e.g., speaking to an unfamiliar peer). The third unit would differ from the first two units in two ways, namely, that the unit's activity levels would (a) be independent of the proximal effects of a specific stimulus and (b) reflect both internal and external processes. With regard to the negative valence domain, subjective reports of anxiety commonly include items that reflect internal processes (e.g., anxious thoughts) and external processes (e.g., behavioral avoidance), and often include items without reference to anxiety-provoking stimuli (De Los Reyes & Makol, 2019; Silverman & Ollendick, 2005). In these respects, including a subjective self-report of trait anxiety or some other context-independent construct (e.g., worry) would facilitate the strategic selection of units that triangulate on a *Trait* score for precisely estimating activity in the target RDoC domain (i.e., negative valence; Figure 2b). Overall, we encourage future research seeking to leverage the models we described for understanding and integrating the RDoC's units of analysis in a way that is not only developmentally sensitive to youth tasks and expectations, but also mindful of proximal and distal contextual influences that may affect results related to these units of analysis.

### **Overview of Special Section Articles and Concluding Comments**

In many respects, we wrote this introductory article from an aspirational perspective. That is, we highlighted measurement considerations that will likely translate to RDoC-informed research producing findings that display low correspondence within and across units of analysis. However, we also provided an overview of a body of research pointing to an innovative idea: low correspondence among units of analysis may *enhance*, not *hinder*, our understanding of youth mental health. In turn, we described models for interpreting and integrating data from

multiple sources, informed by research on understanding subjective reports provided by the multiple informants who commonly contribute data within clinical assessments of youth mental health. We then described how investigators conducting RDoC-informed research might apply these models for enhancing the interpretability of the outcomes of multi-unit assessments. In highlighting these issues, we sought to facilitate the Special Section's meeting both long-term and short-term objectives, described below.

### **Long-Term Objectives of Special Section**

The ideas shared in this article may facilitate addressing long-term objectives regarding RDoC-informed research with youth. In particular, we want the RDoC initiative to avoid past (and present) mistakes in clinical research. We are perhaps most concerned with the prospect of scholars conducting RDoC-informed research, observing low correspondence among units of analysis, and reconciling this low correspondence by omitting some units currently in the framework in favor of retaining a reduced set of “prized” units. Clinical researchers have taken this path before, and we argue, to the detriment of the science and, ultimately, clinical practice.

For instance, decades ago, researchers realized that whether the evidence supports an intervention tested within randomized controlled trials depends on the outcome measure (for reviews, see Casey & Berman, 1985; De Los Reyes & Kazdin, 2006; Weisz, Weiss, Alicke, & Klotz, 1987; Weisz, Weiss, Han, Granger, & Morton, 1995). In response, researchers began adopting a “primary” outcome measure strategy that involves identifying, on an *a priori* basis, a single measure upon which any decisions regarding evidentiary support should be based (for a review, see De Los Reyes, Kundey, & Wang, 2011). Within this strategy, any other indices used to assess outcomes are deemed “secondary” outcome measures. Accordingly, researchers ascribe less “weight” to secondary measures when estimating the effects of the intervention(s) examined.

To be fair, many researchers leverage the primary outcome measure strategy with the best of intentions: to reduce the likelihood that at the conclusion of the study, they “cherry pick” the data that fit their hypotheses (e.g., highlight measures that support the intervention, downplay those that do not). Yet, two factors indicate that this strategy results in *less*, not *more*, knowledge about intervention effects. First, in a significant proportion of peer-reviewed research that uses the primary outcome measure method to report the findings of randomized controlled trials, researchers engage in “spin” tactics in one or more parts of the article (e.g., abstract, results, discussion) to make the intervention effects appear more robust than they actually are (Boutron, Dutton, Ravaud, & Altman, 2010). Stated another way, to the degree that the mission of the primary outcome measure method involves “keeping researchers honest,” the strategy often fails. In fact, the strategy results in practices (e.g., spin tactics) that introduce systematic biases in the empirical process. These biases reduce the likelihood that a given study yields knowledge that informs clinical practices, in that it depresses rates of replication of intervention effects within and across studies (see also Ioannidis, 2005; Open Science Collaboration, 2015).

Second, secondary outcome measures often assess the same domain as the primary outcome measure (De Los Reyes et al., 2011), though previously we reviewed a considerable body of evidence supporting the notion that measures of the same domain might yield discrepant findings for meaningful reasons. In this respect, the primary outcome measure strategy literally results in researchers learning less about the effects of the interventions than alternative strategies that seek to strategically integrate data from multiple measures (e.g., Kraemer et al., 2003; Makol et al., 2020).

Our long-term objective for the Special Section involves *normalizing* the ubiquity of low correspondence in multi-unit assessments conducted within RDoC-informed research. By

highlighting the expected nature of this phenomenon, we will find ourselves in an optimal position to take an empirically informed stance to understanding why this low correspondence exists. In this way, we can avoid common mistakes made in clinical research. Research that reveals instances in which low correspondence signals clinically meaningful information regarding domain-relevant phenomena will pragmatically inform paradigms for strategically integrating data across units of analysis.

### **Short-Term Objectives of Special Section**

In the short-term, we need to start estimating levels of correspondence among units of analysis implemented in RDoC-informed research. Following work that characterizes levels of between-unit correspondence, we need to conduct further research to determine whether variations in levels of correspondence reflect variations in clinically meaningful processes. As a first step, Clarkson and colleagues (2020) leveraged the RDoC initiative's units of analysis framework presented in Table 1 to meta-analyze levels of between-unit correspondence in RDoC-informed research on social processing. This article provides readers with an exemplary illustration of how to estimate these correspondence levels in a developmentally informed manner. Additionally, a pair of commentaries by leading scholars in youth mental health and developmental psychopathology (Beauchaine & Hinshaw, 2020; Garber & Bradshaw, 2020) discuss the research and clinical implications of the Special Section and highlight key directions for future research. Collectively, we expect this Special Section to stimulate thinking about paradigms for optimizing the interpretability of comprehensive, biopsychosocial assessments of youth mental health, as well as inform prevention and treatment models that consider implications of the RDoC initiative's units of analysis.

### 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*, 94-98. doi: 10.1111/j.0963-7214.2006.00414.x
- 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*, 361-382. doi: 10.1037/0033-2909.131.3.361
- 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
- Adrian, M., Zeman, J., & Veits, G. (2011). Methodological implications of the affect revolution: A 35-year review of emotion regulation assessment in children. *Journal of Experimental Child Psychology, 110*, 171-197. doi: 10.1016/j.jecp.2011.03.009
- Akinci, I. H., Tarter, R. E., & Kirisci, L. (2001). Concordance between verbal report and urine screen of recent marijuana use in adolescents. *Addictive Behaviors, 26*, 613-619. doi: 10.1016/S0306-4603(00)00146-5
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders-text revision* (5th ed.) (DSM-5). Washington, DC: Author.
- 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
- Anderson, E. R., Veed, G. J., Inderbitzen-Nolan, H. M., & Hansen, D. J. (2010). An evaluation of the applicability of the tripartite constructs to social anxiety in adolescents. *Journal of*

*Clinical Child and Adolescent Psychology*, 39, 195-207.

doi: 10.1080/15374410903532643

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. doi:

10.1146/annurev-clinpsy-021815-093015

Beauchaine, T.P., & Hinshaw, S.P. (2020). Promises and challenges of the Research Domain Criteria for studying psychopathology among youth. *Journal of Clinical Child and Adolescent Psychology*.

Beauchaine, T. P., Neuhaus, E., Brenner, S. L., & Gatzke-Kopp, L. (2008). Ten good reasons to consider biological processes in prevention and intervention research. *Development and Psychopathology*, 20, 745-774. doi: 10.1017/S0954579408000369

Becker, S.P., Luebke, 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

Beidel, D. S., 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

Blom, E. H., Olsson, E. M., Serlachius, E., Ericson, M., & Ingvar, M. (2010). Heart rate variability (HRV) in adolescent females with anxiety disorders and major depressive disorder. *Acta Paediatrica*, 99, 604-611. doi: 10.1111/j.1651-2227.2009.01657.x

- Boutron, I., Dutton, S., Ravaud, P., & Altman, D. G. (2010). Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *Journal of the American Medical Association, 303*, 2058-2064.  
doi: 10.1001/jama.2010.651
- Bubier, J., Drabick, D. A., & Breiner, T. (2009). Autonomic functioning moderates the relation between contextual factors and externalizing behaviors among inner-city children. *Journal of Family Psychology, 23*, 500-510. doi: 10.1037/a0015555
- Buss, K. A., Davidson, R. J., Kalin, N. H., & Goldsmith, H. H. (2004). Context-specific freezing and associated physiological reactivity as a dysregulated fear response. *Developmental Psychology, 40*, 583-594. doi: 10.1037/0012-1649.40.4.583
- 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*. Advance online publication. doi: 10.1007/s10567-020-00314-4
- Carey, W. B. (1998). Temperament and behavior problems in the classroom. *School Psychology Review, 27*, 522-533. Retrieved from <http://psycnet.apa.org/psycinfo/1999-00154-005>
- 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, R. J., & Berman, J. S. (1985). The outcomes of psychotherapy with children. *Psychological Bulletin, 98*, 388-400. doi: 10.1037/0033-2909.98.2.388
- Cicchetti, D. (1984). The emergence of developmental psychopathology. *Child Development, 55*, 1-7. doi: 10.2307/1129830

- Clarkson, T., Kang, E., Capriola-Hall, N., Lerner, M. D., Jarcho, J., & Prinstein, M. J. (2020). Meta-analysis of the RDoC social processing domain across units of analysis in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*. Advance online publication. doi: 10.1080/15374416.2019.1678167
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10-23. doi: 10.11016/j.brat.2014.04.006
- Crowell, S. E., Beauchaine, T. P., McCauley, E., Smith, C. J., Vasilev, C. A., & Stevens, A. L. (2008). Parent-child interactions, peripheral serotonin, and self-inflicted injury in adolescents. *Journal of Consulting and Clinical Psychology*, 76, 15–21. doi: 10.1037/0022-006X.76.1.15
- Cuthbert, B., & Insel, T. (2010). The data of diagnosis: New approaches to psychiatric classification. *Psychiatry*, 73, 311-314. doi: 10.1521/psyc.2010.73.4.311
- Davis III, T. E., & Ollendick, T. H. (2006). Empirically supported treatments for specific phobia in children: Do efficacious treatments address the components of a phobic response? *Clinical Psychology: Science and Practice*, 12, 144-160. doi: 10.1093/clipsy.bpi018
- Delaney-Black, V., Chiodo, L. M. C., Hannigan, J. H., Greenwald, M. K., Janisse, J., Patterson, G., . . . Sokol, R. J. (2010). Just say ‘I don’t’: Lack of concordance between teen report and biological measures of drug use. *Pediatrics*, 126, 887-893. doi: 10.1542/peds.2009-3059
- De Los Reyes, A. (2011). Introduction to the special section. 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., & 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*, 221-237. doi: 10.1080/15374416.2014.891227

De Los Reyes, A., Aldao, A., Thomas, S.A., Daruwala, S., Swan, A. J., Van Wie, M., ...

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*, 308-322. doi: 10.1007/s10862-012-9289-2

De Los Reyes, A., Augenstein, T. M., & Aldao, A. (2017). Assessment issues in child and adolescent psychotherapy. In J. R. Weisz and A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (3rd ed., pp. 537-554). New York: Guilford.

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., Cook, C. R., Gresham, F. M., Makol, B. A., & Wang, M. (2019). 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. doi: 10.1016/j.jsp.2019.05.005

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., Lerner, M. D., Keeley, L. M., Weber, R., Drabick, D. A. G., Rabinowitz, J., & Goodman, K. L. (2019). Improving interpretability of subjective assessments about psychological phenomena: A review and cross-cultural meta-analysis. *Review of General Psychology*, 23, 293-319. doi: 10.1177/108926801983764
- De Los Reyes, A., & Kazdin, A. E. (2006). Conceptualizing changes in behavior in intervention research: The range of possible changes model. *Psychological Review*, 113, 554-583. doi: 10.1037/0033-295X.113.3.554
- De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review*, 31, 829-838. doi:10.1016/j.cpr.2011.03.011
- De Los Reyes, A., & Makol, B. A. (2019). Evidence-based assessment. In T. H. Ollendick, L. Farrell, and P. Muris (Eds.), *Innovations in CBT for childhood anxiety, OCD, and PTSD: Improving access and outcomes* (pp. 28-51). Cambridge.
- De Los Reyes, A., Ohannessian, C. M., & Racz, S. J. (2019). Discrepancies between adolescent and parent reports about family relationships. *Child Development Perspectives*, 13, 53-58. doi: 10.1111/cdep.12306
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123-149. doi: 10.1146/annurev-clinpsy-050212-185617
- Dembo, R., Briones-Robinson, R., Barrett, K., Winters, K.C., Ungaro, R., Karas, L., ... Wareham, J. (2015). The validity of truant youths' marijuana use and its impact on alcohol use and sexual risk taking. *Journal of Child and Adolescent Substance Abuse*,

24, 355-365. doi: 10.1080/1067828X.2013.844089

- 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
- Deveney, C. M., Grasso, D., Hsu, A., Pine, D. S., Estabrook, C. R., Zobel, E., . . . Briggs-Gowan, M. J. (2019). Multi-method assessment of irritability and differential linkages to neurophysiological indicators of attention allocation to emotional faces in young children. *Developmental Psychobiology*. Advance online publication. doi: 10.1002/dev.21930
- Dishion, T. J., Nelson, S. E., & Bullock, B. M. (2004). Premature adolescent autonomy: parent disengagement and deviant peer process in the amplification of problem behaviour. *Journal of Adolescence, 27*, 515-530. doi: 10.1016/j.adolescence.2004.06.005
- 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., & Kendall, P. C. (2010). Developmental psychopathology and the diagnosis of mental health problems among youth. *Clinical Psychology: Science and Practice, 17*, 272-280. doi: 10.1111/j.1468-2850.2010.01219.x
- 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*, 435-453. doi: 10.1093/clipsy.7.4.435

- Edwards, A. C., Dodge, K. A., Latendresse, S. J., Lansford, J. E., Bates, J. E., Pettit, G. S., Budde, J. P., Goate, A. M., & Dick, D. M. (2011). MAOA uVNTR and early physical discipline interact to influence delinquent behavior. *Journal of Child Psychology and Psychiatry, 51*, 679-687. doi: 10.1111/j.1469-7610.2009.02196.x
- Garber, J., & Bradshaw, C.P. (2020). Optimizing the RDoC model through developmental adaptations: Increasing the potential for transdiagnostic utility. *Journal of Clinical Child and Adolescent Psychology*.
- Gillissen, R., Bakermans-Kranenburg, M. J., van Ijzendoorn, M. H., & Linting, M. (2008). Electrodermal reactivity during the Trier Social Stress Test for children: Interaction between the serotonin transporter polymorphism and children's attachment representation. *Developmental Psychobiology, 50*, 615-625. doi: 10.1002/dev.20314
- 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
- Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review, 106*, 458-490. doi: 10.1037/0033-295X.106.3.458
- Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*, 101-131. doi: 10.1037/0033-295X.113.1.101

- Greaves-Lord, K., Ferdinand, R. F., Sondejker, F. E. P. L., Dietrich, A., Oldehinkel, A. J., Rosmalen, J. G. M., ... Verhulst, F. C. (2007). Testing the tripartite model in young adolescents: Is hyperarousal specific for anxiety and not depression? *Journal of Affective Disorders, 102*, 55-63. doi: 10.1016/j.jad.2006.12.009
- Groth-Marnat, G., & Wright, A.J. (2016). *Handbook of psychological assessment* (6th ed.). Hoboken, NJ: Wiley.
- Gunnar, M. R., Talge, N. M., & Herrera, A. (2009). Stressor paradigms in developmental studies: What does and does not work to produce mean increases in salivary cortisol. *Psychoneuroendocrinology, 34*, 953-967. doi: 10.1016/j.psyneuen.2009.02.010
- Harris, K. M., Griffin, B. A., McCaffrey, D. F., & Morral, A. R. (2008). Inconsistencies in self-reported drug use by adolescents in substance abuse treatment: Implications for outcome and performance measurements. *Journal of Substance Abuse Treatment, 34*, 347-355. doi: 10.1016/j.jsat.2007.05.004
- Henrich, J., Heine, S., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*, 29. doi: 10.1038/466029a
- Henry, D. B., & Study Research Group, T. M. A. C. (2006). Associations between peer nominations, teacher ratings, self-reports, and observations of malicious and disruptive behavior. *Assessment, 13*(3), 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. doi: 10.1207/s15374424jccp2102\_6
- 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
- Hubbard, J. A., Parker, E. H., Ramsden, S. R., Flanagan, K. D., Relyea, N., Dearing, K. F., . . . Hyde, C. T. (2004). The relations among observational, physiological, and self-report measures of children's anger. *Social Development*, 13, 14-39.  
doi: 10.1111/j.1467-9507.2004.00255.x
- Hunsley, J., & Lee, C. M. (2014). *Introduction to clinical psychology* (2<sup>nd</sup> ed.). Hoboken, NJ, US: Wiley.
- 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
- Hymel, S., Rubin, K. H., Rowden, L., & LeMare, L. (1990). Children's peer relationships: Longitudinal prediction of internalizing and externalizing problems from middle to late childhood. *Child Development*, 61, 2004-2021. doi: 10.1111/j.1467-8624.1990.tb03582.x
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2, e124. doi:10.1371/journal.pmed.0020124
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167, 748-751.  
doi: 10.1176/appi.ajp.2010.09091379
- Kazdin, A. E., & Blasé, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science*, 6, 21-37.  
doi:10.1177/1745691610393527
- Kazdin, A. E., & Kagan, J. (1994). Models of dysfunction in developmental psychopathology.

*Clinical Psychology: Science and Practice*, 1, 35-52. doi: 10.1111/j.1468-2850.1994.tb00005.x

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

Lapouse, R., & Monk, M. A. (1958). An epidemiologic study of behavior characteristics in children. *American Journal of Public Health*, 48, 1134-1144. doi: 10.2105/AJPH.48.9.1134

Lennox, R., Dennis, M. L., Ives, M., & White, M. K. (2006). The construct and predictive validity of different approaches to combining urine and self-reported drug use measures among older adolescents after substance abuse treatment. *The American Journal on Addictions*, 15, 92-101. doi: 10.1080/10550490601006089

Leventon, J. S., Stevens, J. S., & Bauer, P. J. (2014). Development in the neurophysiology of emotion processing and memory in school-age children. *Developmental Cognitive Neuroscience*, 10, 21-33. doi: 10.1016/j.dcn.2014.07.007

Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S., & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *American Journal of Psychiatry*, 165, 124-132. doi: 10.1176/appi.ajp.2007.06091513

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*. Manuscript

submitted for publication.

- Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology, 34*, 362-379. doi: 10.1207/s15374424jccp3403\_1
- Michels, N., Sioen, I., Clays, E., De Buyzere, M., Ahrens, W., Huybrechts, I., . . . De Henauw, S. (2013). Children's heart rate variability as stress indicator: Association with reported stress and cortisol. *Biological Psychology, 94*, 433-440.  
doi: 10.1016/j.biopsycho.2013.08.005
- Miers, A. C., Blöte, A. W., Sumter, S. R., Kallen, V. L., & Westenberg, P. M. (2011). Subjective and objective arousal correspondence and the role of self-monitoring processes in high and low socially anxious youth. *Journal of Experimental Psychopathology, 2*, 531-550. doi: 10.5127/jep.019411
- Miller, A. L., Seifer, R., Stroud, L., Sheinkopf, S. J., & Dickstein, S. (2006). Biobehavioral indices of emotion regulation relate to school attitudes, motivation, and behavior problems in a low-income preschool sample. *Annals of the New York Academy of Sciences, 1094*, 325-329. doi: 10.1196/annals.1376.043
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246-268. doi: 10.1037/0033-295X.102.2.246
- National Institute of Mental Health (2020). *Units of analysis*. Retrieved from <https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/units/index.shtml>
- Ollendick, T. H., & Hirshfeld-Becker, D. R. (2002). The developmental psychopathology of social anxiety disorder. *Biological Psychiatry, 51*, 44-58. doi: 10.1016/S0006-

3223(01)01305-1

- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716. doi:10.1126/science.aac4716
- Raggi, V. L., Samson, J. G., Felton, J. W., Loffredo, H. R., & Berghorst, L. H. (2018). *Exposure therapy for treating anxiety in children and adolescents: A comprehensive guide*. Oakland, CA: New Harbinger Publications.
- Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin*, *112*, 485-499. doi: 10.1037/0033-2909.112.3.485
- Rinnewitz, L., Parzer, P., Koenig, J., Bertsch, K., Brunner, R., Resch, F., & Kaess, M. (2019). A biobehavioral validation of the Taylor Aggression Paradigm in female adolescents. *Nature: Scientific Reports*, *9*, 1-10. doi: 10.1038/s41598-019-43456-4
- Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heinsse, R. K., & ... Cuthbert, B. N. (2010). Developing constructs for psychopathology research: Research domain criteria. *Journal of Abnormal Psychology*, *119*, 631-639. doi: 10.1037/a0020909
- Schmitz, J., Blechert, J., Krämer, M., Asbrand, J., & Tuschen-Caffier, B. (2012). Biased perception and interpretation of bodily anxiety symptoms in childhood social anxiety. *Journal of Clinical Child and Adolescent Psychology*, *41*, 92-102. doi: 10.1080/15374416.2012.632349
- Schmitz, J., Tuschen-Caffier, B., Wilhelm, F. H., & Blechert, J. (2013). Taking a closer look: Autonomic dysregulation in socially anxious children. *European Child and Adolescent Psychiatry*, *22*, 631-640. doi: 10.1007/s00787-013-0405-y
- 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
- Skinner, B. F. (1953). *Science and Human Behavior*. New York: MacMillan.
- 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, 1026-1041. doi: 10.1007/s10803-014-2258-8
- van Doorn, M. M., Bodden, D., Jansen, M., Rapee, R. M., & Granic, I. (2018). Linking mother-child discrepancies to behavioral observations of children's anxiety. *Child and Youth Care Forum*, 47, 481-498. doi: 10.1007/s10566-018-9441-6
- Vanderhasselt, M. A., De Raedt, R., Baeken, C., Leyman, L., Clerinx, P., & D'haenen, H. (2007). The influence of rTMS over the right dorsolateral prefrontal cortex on top-down attentional processes. *Brain Research*, 1137, 111-116. doi: 10.1016/j.brainres.2006.12.050
- Vanderhasselt, M. A., De Raedt, R., Baeken, C., Leyman, L., & D'haenen, H. (2006). The influence of rTMS over the right dorsolateral prefrontal cortex on intentional set switching. *Experimental Brain Research*, 172, 561-565. doi: 10.1007/s00221-006-0540-5
- Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, 40, 472-481. doi:10.1016/j.jrp.2005.03.003
- 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 and Adolescent Psychiatry*, 47, 632-641. doi: 10.1097/CHI.0b013e31816c5c10

- Weisz, J. R., Hawley, K. M., & Jensen Doss, A. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America*, *13*, 729-815. doi: 10.1016/j.chc.2004.05.006
- Weisz, J. R., Jensen Doss, A., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology*, *56*, 337-363. doi: 10.1146/annurev.psych.55.090902.141449
- Weisz, J. R., & Kazdin, A. E. (Eds.). (2017). *Evidence-based psychotherapies for children and adolescents* (3rd ed.). New York: Guilford Press.
- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, *55*, 542-549. doi: 10.1037/0022-006X.55.4.542
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, *117*, 450-468. doi: 10.1037/0033-2909.117.3.450
- Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the internet. *Journal of Personality and Social Psychology*, *79*, 748-762. doi: 10.1037/0022-3514.79.5.748
- Williams, R. J., & Nowatzki, N. (2005). Validity of adolescent self-report of substance use. *Substance Use and Misuse*, *40*, 299-311. doi: 10.1081/JA-200049327
- 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

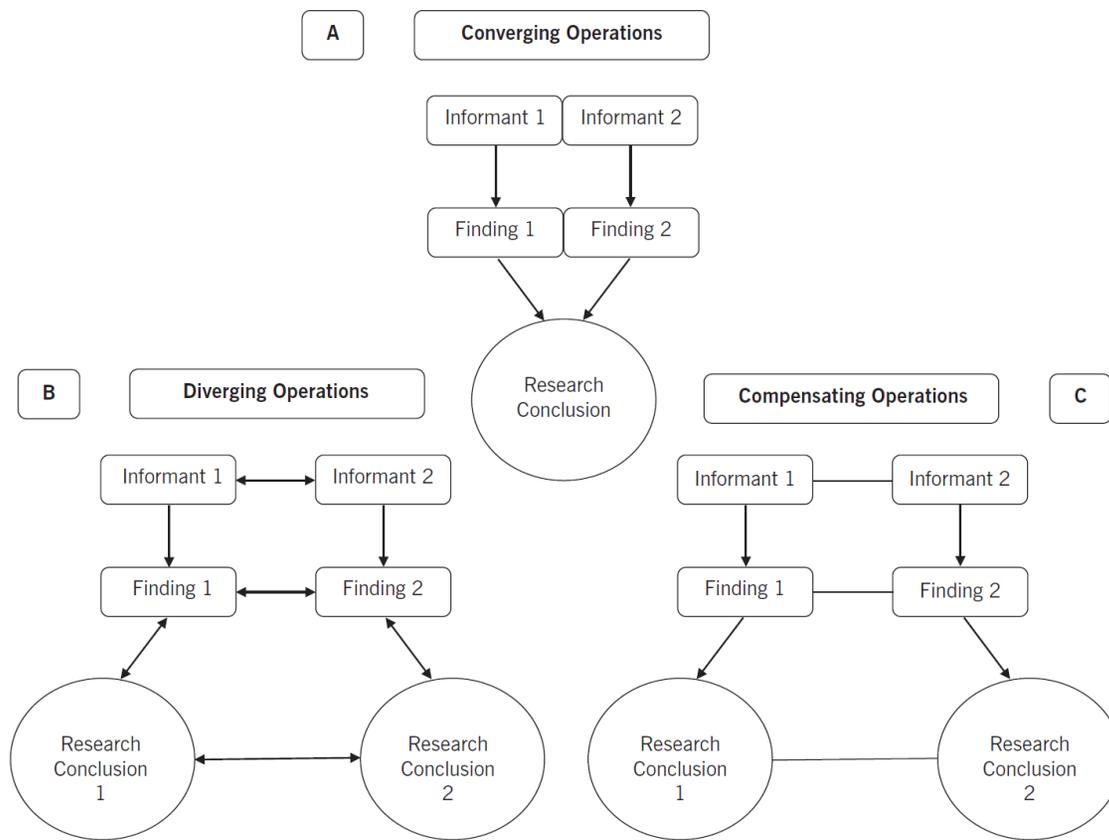
World Health Organization. (2007). *International statistical classification of diseases and related health problems* (10th rev.). Retrieved from <http://apps.who.int/classifications/apps/icd/icd10online>

**Table 1. Research Domain Criteria Units of Analysis Framework: Modification for Use in Research with Youth**

<b>Subjective</b>	Self-Reports
	Adult Caregivers in Home
	Adult Caregivers in School
	Clinical Interviewers (e.g., global assessments of functional impairment; clinical severity ratings)
	Peer Reports
<b>Behavior</b>	Controlled laboratory observations (home-based functioning)
	Controlled laboratory observations (school-based functioning)
	Controlled laboratory observations (social/peer functioning)
	Naturalistic observations at home
	Naturalistic observations at school
	Naturalistic observations with peers/social settings
	Performance-based tasks (e.g., eye tracking and reaction time in response to standardized tasks that assess attention)
	Official records (e.g., school records, legal records)
<b>Physiology</b>	Physiological markers of arousal (e.g., heart rate)
	Physiological markers of behavioral activity (e.g., actigraphy-daytime)
	Physiological markers of sleep functioning (e.g., actigraphy-nighttime; polysomnography)
	Physiological markers of immune system functioning (e.g., c reactive proteins, cytokines)
	Physiological markers of emotional regulation (e.g., heart rate variability; vagal tone, respiratory sinus arrhythmia)
	Physiological markers of motivation/engagement (e.g., pupilometer; eye tracking)
	Physiological markers of threat appraisal (e.g., startle response)
	Physiological markers of stress reactivity (e.g., hormonal assays of HPA axis markers)
	Physiological markers of social behavior/affiliation (e.g., hormonal assays of oxytocin)
	Resting EEG (e.g., asymmetry)
	Brain potentials (e.g., EEG/ERP)
<b>Circuits</b>	e.g., BOLD (activation of cortical signals)
	e.g., BOLD (activation of sub-cortical signals)
	e.g., BOLD (connectivity between/among cortical regions)
	e.g., BOLD (connectivity between/among sub-cortical signals)
	e.g., BOLD (connectivity between/among cortical and sub-cortical regions)
	Resting state fMRI (e.g., default-mode network development)
	Grey/white matter volume
	Diffusion Tensor Imaging (e.g., white matter tracking)
<b>Cells</b>	Cellular assays
<b>Molecules</b>	Molecular assays

<b>Genes</b>	Genotype assays: SNP, DNR, VNTR
	Genotype assays: Allelic risk score
	Genotype assays: GWAS analysis
	Gene x Environment interactions (e.g., as studied in diathesis-stress research, or for assessing differential susceptibility to environments)
<b>Paradigms</b>	<b>A single experimental/clinical paradigm may induce activity along the five RDoC domains. For instance, RDoC studies using the Trier Social Stress Test or Mood Induction Paradigms might induce RDoC domain activity in the valence systems, social processing, and/or cognition domains.</b>

*Note.* BOLD = blood-oxygen-level-dependent; DNR = dinucleotide repeat polymorphism; EEG = electroencephalography; ERP = event-related potentials; fMRI = functional magnetic resonance imaging; GWAS = genome-wide association study; HPA = hypothalamic-pituitary-adrenal; SNP = single nucleotide polymorphism; VNTR = variable number tandem repeat polymorphism



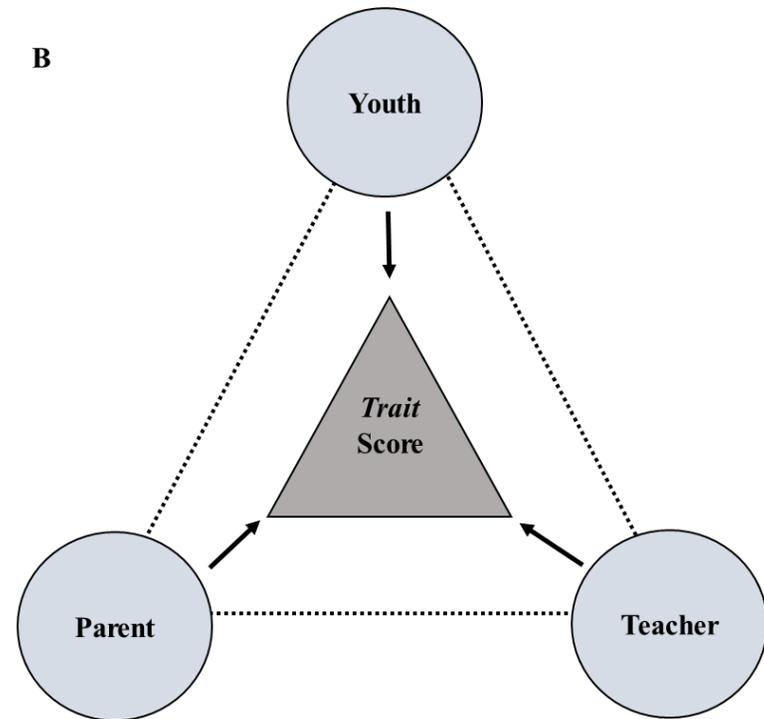
*Figure 1.* Graphical representation of the research concepts that comprise the Operations Triad Model. The top half (A) represents Converging Operations: a set of measurement conditions for interpreting patterns of findings based on the consistency within which findings yield similar conclusions. The 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). © Annual Review of Clinical Psychology. Copyright 2012 Annual Reviews. All rights reserved. The Annual Reviews logo, and other Annual Reviews products referenced herein are either registered trademarks or trademarks of Annual Reviews. All other marks are the property of their respective owner and/or licensor.

A

		<u>Context</u>	
		School	Home
<u>Perspective</u>	Self-Report	Youth	Youth
	Other-Report	Teacher	Parent

B



*Figure 2. 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).

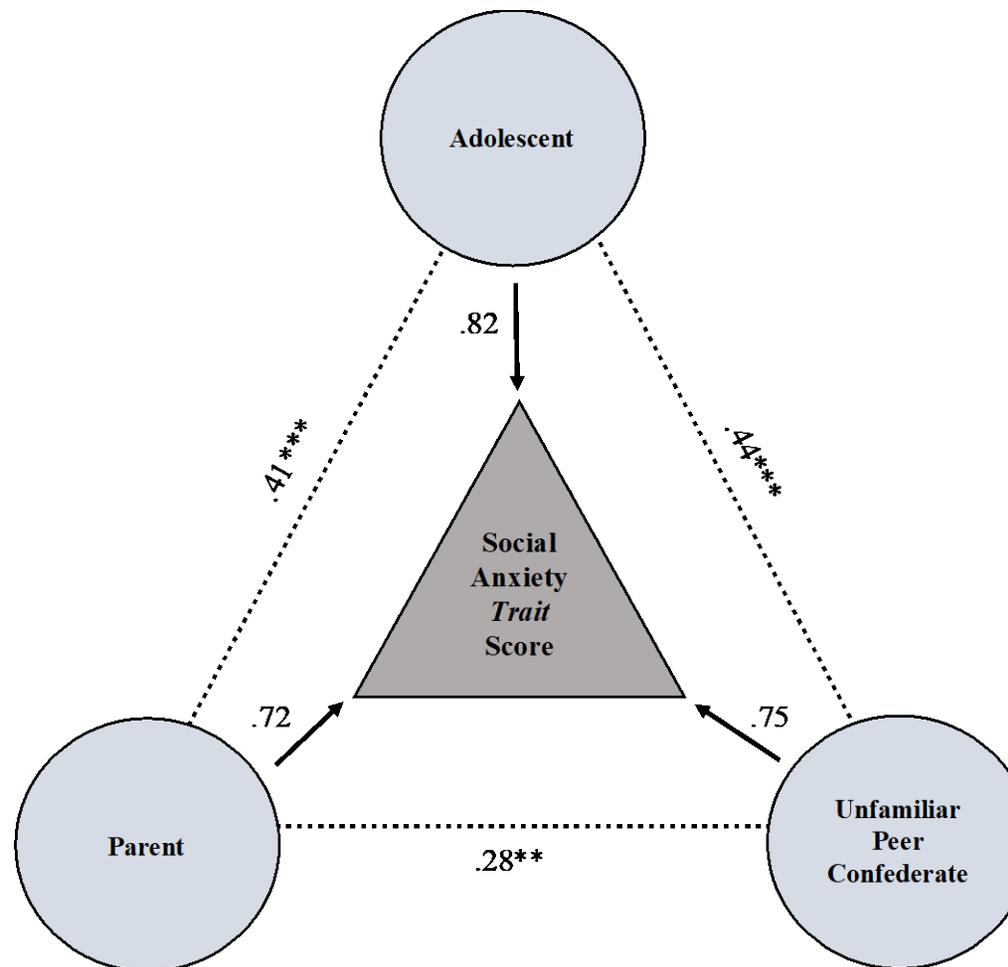
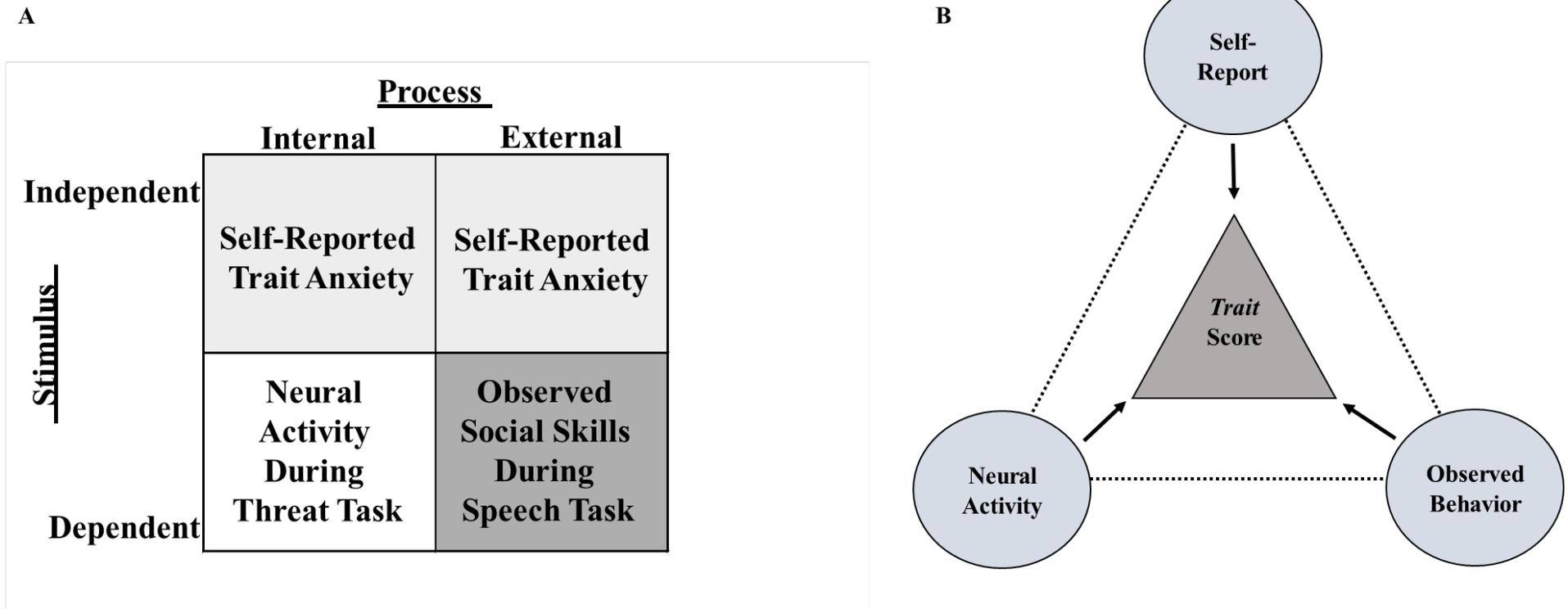


Figure 3. From Makol and colleagues (2020): Values on dotted lines denote bivariate correlations among parent, adolescent, and unfamiliar peer confederate reports of adolescent social anxiety. Values on solid arrows denote component weights from principal components analysis of parent, adolescent, and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



*Figure 4. Panel A:* Example of use of “mix-and-match” criterion to strategically integrate units of analysis in a multi-unit assessment used to index activity relevant to the negative valence RDoC domain. Units systematically vary with regard to whether they assess a *process* for which activity is directly observable or internally expressed, and whether unit activity depends on a stimulus, with the goal of effectively triangulating on a *Trait* score of unit activity. *Panel B:* Graphical depiction of multi-unit activity triangulating, much like GPS, to identify a negative valence *Trait* score. Both neural activity and observed behavior are units for which their activity requires a stimulus (i.e., threat-based cues), with neural activity indexing internal processes and observed behavior indexing external processes. The self-report of trait anxiety estimates activity independent of direct activation of a state-based stimulus, and processes reflecting trait anxiety manifest internally (e.g., anxious thoughts) and externally (e.g., avoidance, pressured speech). Figures adapted from Kraemer et al. (2003).