Running Head: HiTOP RESEARCHER'S PRIMER

A Hierarchical Taxonomy of Psychopathology (HiTOP) Primer for Mental Health Researchers

Christopher C. Conway[1], Miriam K. Forbes[2†], Susan C. South [3†], & the HiTOP Consortium

[1] Department of Psychology, Fordham University, Bronx, NY, USA

[2] Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney, Australia

[3] Purdue University, Department of Psychological Sciences, West Lafayette, IN, USA

[†] Joint second authors, listed in alphabetical order

Corresponding author

Christopher C. Conway, Department of Psychology, Fordham University, 424 Dealy Hall, 441 E Fordham Road, Bronx, NY 10458, USA; email: cconway26@fordham.edu

Author Note

This article was organized by the HiTOP Consortium (https://medicine.stonybrookmedicine.edu/HITOP).

Author Contributions

C.C.C. formulated the paper concept. C.C.C., M.K.F., and S.C.S. wrote the paper. A.J.S., C.C.C., and M.K.F created figures. M.K.F. and S.C.S. are joint second authors. The following authors (listed alphabetically) contributed to reviewing and revising the paper: Marina Bornovalova, Raymond Chan, Michael Chmielewski, Lee Anna Clark, Tim Dalgleish, Danielle Dick, Michael Dretsch, Nicholas Eaton, Alexander Fornito, Vina Goghari, John D. Haltigan, Benjamin Hankin, Christopher Hopwood, Katherine Jonas, Roman Kotov, Robert Krueger, Robert Latzman, Donald Lyman, Elizabeth Martin, Giorgia Michelini, Joshua Miller, Terrie Moffitt, Stephanie Mullins-Sweatt, Kristin Naragon-Gainey, Thomas Olino, Christopher Patrick, Aaron Pincus, Craig Rodriguez-Seijas, Douglas Samuel, Martin Sellbom, Alexander J. Shackman, Kasey Stanton, Jeggan Tiego, Irwin Waldman, Monika Waszczuk, David Watson, Ashley Watts, Mark Waugh, Sylia Wilson, Aidan Wright, Jami Young, and David Zald.

Resource Sharing

Data and annotated code for the tutorial examples are freely available at the Open Science Framework website (https://osf.io/8myzw).

Abstract

Mental health research is at an important crossroads as the field seeks more reliable and valid phenotypes to study. Dimensional approaches to quantifying mental illness operate outside the confines of traditional categorical diagnoses, and they are gaining traction as a way to advance research on the causes and consequences of mental illness. The Hierarchical Taxonomy of Psychopathology (HiTOP) is a leading dimensional research paradigm that synthesizes decades of investigation into evidence-based dimensions of psychological disorders. This article aims to make the application of dimensional approaches in mental health research more accessible through tutorials that demonstrate how to use the HiTOP model to formulate and test research questions. Data and annotated code (written for R and Mplus software programs) for each example are included (https://osf.io/8myzw). We outline how investigators can use these ideas and tools to generate new insights in their own substantive research programs.

Keywords: assessment, classification, Hierarchical Taxonomy of Psychopathology (HiTOP), nosology, psychopathology, transdiagnostic

A Hierarchical Taxonomy of Psychopathology (HiTOP) Primer for Mental Health Researchers

Mental health research is at a crossroads. Historically the field has relied on categorical diagnoses as the basic units of analysis, but there is mounting concern about the validity of these categories, as currently defined. Many believe that diagnostic categories are limiting much-needed insights into disorder etiology, treatment, and prevention (e.g., Gordon & Redish, 2016). This criticism has put the focus on alternative *dimensional* perspectives that prioritize phenotypes that cut across traditional diagnostic boundaries (Kotov et al., 2017; Kozak & Cuthbert, 2016).

The Hierarchical Taxonomy of Psychopathology (HiTOP) is an empirically derived model of the major dimensions of mental illness. It represents an alternative research paradigm that, as we argue below, has multiple advantages relative to categorical rubrics. Several publications describe HiTOP's conceptual and empirical foundations (Kotov et al., 2017, Krueger et al., 2018), but as yet there are no resources that explain how mental health researchers can design and interpret results from studies framed by the HiTOP system[1]. In this article, we provide tutorials and empirical examples for prototypical research questions that can be formulated and tested in the HiTOP framework. We also consider obstacles and limitations to this approach, as well as ongoing methodological developments. Our aim is to make it clearer how the HiTOP model can be applied in mental health research across a variety of substantive areas, and to give an overview of the relevant statistical methods. We include the data and

[1] We intend for these tutorials to be self-contained, but researchers who want to learn more about HiTOP's origins and applications can refer to other materials published by the HiTOP consortium (e.g., Conway et al., 2019; Kotov et al., 2017; Krueger et al., 2018; Ruggero et al., 2019, Waszczuk et al., 2020). Several scientific journals have recently printed special issues dedicated to examining HiTOP's philosophy and research applications (Conway & Simms, 2020; Eaton, 2017; Krueger, Tackett, & MacDonald, 2016; Krueger, Watson, & Widiger, 2020). Support and information also are available from the consortium's website.

annotated code for each of our examples in the supplemental material, and cite additional references for more detailed technical information on the statistical methods throughout.

Hierarchical Taxonomy of Psychopathology (HiTOP)

To put the tutorials in context, we begin with a brief review of the motivation behind and makeup of the HiTOP model of psychopathology. We refer readers to prior consortium publications for the full account of HiTOP's dimensional structure (Kotov et al., 2017; Krueger et al., 2018) and its implications for substantive research on the origins, development, consequences, and treatment of psychopathology (Conway et al., 2019; Latzman, DeYoung, & The HiTOP Neurobiological Foundations Workgroup, 2020; Ruggero et al., 2019; Waszczuk et al., 2020).

The HiTOP consortium formulated a new nosological system in response to problems with traditional categorical nosologies like the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; American Psychiatric Association [APA], 2013) and the *International Classification of Diseases* (*ICD*; World Health Organization, 2018). Issues like rampant comorbidity, excessive within-diagnosis heterogeneity, and unreliability put significant constraints on the utility of categorical diagnoses for assessment and applied research (Clark, Cuthbert, Lewis-Fernández, Narrow, & Reed, 2017; Widiger & Samuel, 2005).

Figure 1 illustrates the current working model for HiTOP. The upper part of the model is oriented around *spectrum* constructs: internalizing, antagonistic externalizing, disinhibited externalizing, detachment, thought disorder, and somatoform. The first five spectra are particularly well-validated, and are akin to dimensions described in consensus models of both adaptive and maladaptive personality (Markon et al., 2005), whereas somatoform has more modest empirical backing (e.g., Michelini et al., 2019). The shared characteristics of the six

HiTOP spectra are captured in a *superspectrum*, the general factor of psychopathology (dubbed the *p*-factor by Caspi et al., 2014), which is presumed to pervade all mental disorder symptoms (albeit some more than others) (Lahey et al., 2012). Moving down the hierarchy, the six broad dimensions divide into more homogeneous dimensions that are labeled *subfactors*. For instance, internalizing is subdivided into fear and distress, among others. These subfactors are, in turn, composed of lower-order *syndromes*, which are intended to be empirically derived symptom constellations. In practice, our current understanding of the lower levels of the hierarchy is largely bound to DSM constructs, in line with Figure 1, but this area of the literature is evolving rapidly, and the empirically derived symptom constellations may not directly correspond to the syndrome concepts in the DSM as the HiTOP framework evolves (see, e.g., Waszczuk, Kotov, Ruggero, Gamez, & Watson, 2017). At the base of the hierarchy, *symptom components* and *maladaptive personality traits* represent the basic units of psychological disorders. We use terms like *superordinate* and *higher-order* to refer to broader dimensions in the model and terms like *subordinate* and *lower-order* for narrower dimensions.

HiTOP as a Research Framework

The HiTOP model has a descriptive function; it delineates the dimensional phenotypes that characterize mental health conditions. At the same time, the model has clear implications for how psychological theories are formulated and tested, and for how research is designed and interpreted. HiTOP dimensions change not only the way we ask questions about the causes, correlates, and consequences of mental illness, but also the very questions that we ask.

Dimensions

The HiTOP model reflects a paradigmatic break from traditional classification schemes, which assume that someone either has or does not have a condition, with no middle ground.

Instead, the HiTOP model conceptualizes psychopathology in terms of continuously distributed dimensions, an approach that is in sync with a great deal of research evidence (Haslam et al., 2020). This has implications for the sort of assessment information we use to test hypotheses about mental disorders. Specifically, the focus is on dimensional representations of mental health problems, as measured directly or abstracted via latent variable modeling from observable signs and self-reported symptoms of mental disorder.

Hierarchy

HiTOP assumes that the building blocks of psychopathology are related to one another hierarchically, paralleling the configuration of other individual difference domains, such as personality, intelligence, and executive functioning (e.g., Deary, 2012; Markon, Krueger, & Watson, 2005). This conceptualization means that mental health conditions can be construed at various levels of breadth. Consider social anxiety disorder: Using Figure 1, researchers can operationally define social anxiety in terms of (a) fine-grain symptom components, such as situational avoidance; (b) a cluster of symptom components that reflect common manifestations of social anxiety; (c) fear responses it shares with panic and other phobic disorders; (d) emotional disturbances common to all internalizing problems; and, at the broadest level, (e) general distress and impairment, reflected in the *p*-factor, which saturates almost all psychopathology. The different levels are neither incommensurate nor mutually exclusive; they are equally valid representations of social anxiety in the HiTOP system, albeit differing in specificity; just as, for example, in biological classification animals can have various taxonomic ranks (e.g., for dogs: kingdom = animalia, phylum = chordata, class = mammalia, and so on) .

This perspective enables researchers to pinpoint the "level of analysis" that conveys the most predictive information for a given research context. Stated differently, they can identify

what part(s) of a mental health condition are associated with outcomes of interest. This versatility opens up entirely new research questions and hypotheses. For example, is the link between posttraumatic stress disorder and romantic relationship difficulties best explained by its associations with (a) fairly specific features, such as hyperarousal and hostility, are associated with romantic difficulties (symptom component level); (b) many or all of its component parts are associated with romantic difficulties (syndrome level); or (c) emotional dysfunction common to all anxiety and depressive conditions (spectrum level) is associated with romantic difficulties? Research framed by *DSM* categories cannot empirically compare these hypotheses because the categories collapse individual differences across these various levels of a dimensional hierarchy into a single level of analysis (i.e., the binary syndrome). Research informed by HiTOP can shed new light on the nature of connections between psychological symptoms and their hypothesized causes and consequences.

Measurement

What does HiTOP measurement look like, practically speaking? How do hierarchically arrayed dimensions make their way into your dataset? These are natural questions, because many of the terms in Figure 1 are less familiar than the *DSM-5* labels to many researchers. Keep in mind, though, that HiTOP constructs comprise the same clinical phenomena that researchers and clinicians work with routinely—they are simply reorganized based on structural empirical evidence.

HiTOP constructs can be (1) accessed directly with existing measurement instruments and also (2) captured in the shared variance among scores on traditional measures using latent variable modeling. We now discuss each of these options in turn.

Direct assessment. Many of the constructs described in Figure 1 can be assessed with surveys and interviews that were developed with the hierarchical structure of psychopathology in mind. One advantage of this option is that these symptom scales were created after verifying certain psychometric desiderata, such as reliability and discriminant validity, in normative samples. These features are not assured when working with symptom-level data derived from assessments oriented around categorical diagnoses.

Several such empirically derived measures already exist. For instance, the Inventory of Depressive and Anxiety Symptoms (IDAS) was designed to measure the core components of the internalizing spectrum (Watson et al., 2007). Framed around the influential tripartite theory of anxiety and depression (and its successors) (Clark & Watson, 1991; Mineka, Watson, & Clark,1998; Watson, 2009), the IDAS features empirically derived subscales that tap into homogeneous components of emotional problems, such as panic, insomnia, and social anxiety. A revised IDAS, the IDAS-II, includes additional scales related to posttraumatic stress, obsessive-compulsive, and bipolar disorders (Watson et al., 2012). The Externalizing Spectrum Inventory is the analog for the externalizing domain (Krueger, Markon, Patrick, Benning, & Kramer, 2007). It contains 23 narrow-bandwidth facet scales (e.g., alcohol use, theft, fraud) that cohere empirically into 3 broader dimensions (viz., general disinhibition, callous aggression, and substance abuse). The Schedule for Nonadaptive and Adaptive Personality (Clark, Simms, Wu, & Casillas, 2014), Five Factor Model personality disorder scales (Widiger, Lynam, Miller, & Oltmanns, 2012), and 220- and 100-item versions of the PID-5 (Krueger et al., 2012; Maples et al., 2015) all include various lower-order scales that capture maladaptive personality traits characterizing clinical and personality disorders, as well as broader trait concepts that reflect the major dimensions of Big Three and Big Five models of personality (e.g., Markon et al., 2005).

Also, some broadband clinical inventories (e.g., MMPI-2-RF [Ben-Porath & Tellegen, 2008], Personality Assessment Inventory [Morey, 2007]; SPECTRA: Indices of Psychopathology [Blais & Sinclair, 2018]) have scales that structurally adhere broadly to contemporary psychopathology frameworks (e.g., Lee, Sellbom, & Hopwood, 2017).

Finally, the HiTOP consortium is now in the process of developing a measure of the higher- and lower-order dimensions in 6 HiTOP domains: internalizing, detachment, thought disorder, somatoform pathology, and both disinhibited and antagonistic forms of externalizing psychopathology (Simms et al., 2020). After an omnibus self-report questionnaire is finalized, there are plans also to create informant-report and interview measures.

Latent variable modeling. Factor analysis and structural equation modeling are families of statistical approaches that naturally align with hierarchical dimensional models. Latent variables can be used to model the higher-order dimensions in the HiTOP framework (e.g., subfactors, spectra, and superspectra) by capturing the shared variance among psychopathology constructs in a dataset. For example, in a study that assesses mental disorder diagnoses using a structured clinical interview, an internalizing latent variable could be modeled based on the patterns of comorbidity (i.e., covariance) among major depression, generalized anxiety, social anxiety, and panic disorders. Similarly, in a study that assesses a variety of self-reported symptoms of psychopathology using the Strengths and Difficulties Questionnaire (Goodman, 2001) or Youth Self-Report (Achenbach, 2009), broader internalizing and externalizing spectra can be modeled as correlated latent variables based on the overlap among the corresponding scales. For interested readers who have not used these techniques before, there are a number of gentle introductions to factor analysis and structural equation modeling available that cover the

assumptions, specification, and evaluation of latent variable models (e.g., Brown, 2015; Kline, 2015).

More complex latent variable models can also be used to operationalize multiple levels of the HiTOP framework simultaneously. For example, the literature on the nature and correlates of general psychopathology (or "*p*-factor") suggests that this dimension embodies individuals' propensity to all forms of psychopathology (e.g., Caspi & Moffitt, 2018). As mentioned earlier, it rests at the apex of the HiTOP model and has been hypothesized to capture the non-specific etiological factors that increase risk for all symptom domains (Lahey et al., 2017).

Bifactor modeling is a popular approach for representing the general factor (e.g., Lahey et al., 2012; Caspi et al., 2014). In these latent variable models, all of the observed variables—for example, categorical diagnoses, or continuous scales of different symptom domains—load on a single general factor that captures their shared variance (general psychopathology), and the remaining variance is partitioned into uncorrelated, narrower specific factors (e.g., internalizing, externalizing, and thought disorder). Multiple limitations of the bifactor model have been highlighted in recent empirical and simulation research (e.g., Greene et al., 2019; Watts, Poore, & Waldman, 2019), leading to an increasing number of studies using a higher-order model where a second-order latent variable (general psychopathology) is modeled atop several first-order latent variables (internalizing, externalizing, and thought disorder) to capture the correlations among them (see Tutorial 2).

Although covering the relative strengths and weaknesses of these approaches is beyond the scope of this paper, we direct interested readers to a recent study that comprehensively compared the most popular models used in research on the structure of psychopathology (Forbes et al., in press; see also Bornovolova et al., 2020). Comparing the reliability and validity of the

latent variable models commonly used in research using a HiTOP framework, Forbes et al. found that the general factors of psychopathology in the bifactor and higher-order models were both robust and performed similarly well. The results of the study boiled down to three core recommendations to researchers using these models to understand the structure and correlates of HiTOP constructs: (1) Look beyond traditional model fit indices to choose which model fits "best" for the data and research question; (2) examine the reliability of latent variables directly; and (3) be cautious when isolating and interpreting the unique effects of individual latent variables (see further discussion in Tutorial 1 and footnote 3).

Common Research Designs

There are several ways that the HiTOP system tends to be applied in substantive research (see Conway et al., 2019, for a review). First, investigators have used latent variable modeling techniques to establish the number and nature of latent dimensions relevant to a psychopathology domain (e.g., Forbes et al., 2020; Michelini et al., 2019). Often called "structural modeling," this line of inquiry resolves the dimensional phenotypes that might serve as research or assessment targets. Thus, we can consider it as a precursor to tests of psychopathology's association with putative causes and outcomes.

Second, researchers study how the heterogeneous symptom components of a diagnosis (or small cluster of diagnoses) differentially relate to a given outcome. This approach elucidates the fine-grain clinical problems that contribute to an outcome of interest. Consider how major depression can be decomposed into affective, vegetative, and cognitive symptom components. Researchers can simultaneously evaluate the predictive validity of these dimensions in relation to clinically important outcomes (e.g., emergency room visits).

Third, investigators compare the criterion validity of dimensions at varying levels of breadth. The psychopathology phenotypes in these studies span multiple levels of the HiTOP hierarchy. The objective is to determine whether one level of the dimensional hierarchy (e.g., spectrum), compared to another (e.g., symptom component), better predicts important outcomes.

Tutorials

In this section we present three tutorials for mental health researchers. Tutorial 1 illustrates how to partition a broad psychopathology phenotype into relatively narrow individual differences and, in turn, how to compare their prediction of an outcome of interest. Tutorial 2 derives a hierarchy of dimensions of emotional disorders from a battery of symptom-level data and then examines phenotypes across multiple levels in relation to romantic relationship quality. Tutorial 3 covers how to integrate higher-order dimensions of mental disorder into developmental psychopathology designs.

Data, codebooks, and analysis code for all three tutorials covered in this manuscript are available at https://osf.io/8myzw. We supply annotated code written for both R and Mplus software platforms.

Tutorial 1: Comparing Specific Symptom Components' Effects

Earlier we alluded to the fact that diagnostic heterogeneity complicates mental health research framed around binary disorders. When diverse clinical problems can contribute to a diagnosis, two people with the same condition can exhibit very different symptom profiles (e.g., Fried & Nesse, 2015). This within-diagnosis variability is problematic for applied researchers because it is difficult to know which components of a disorder explain its correlation with a particular risk factor or outcome.

The HiTOP framework allows more control over the issue of heterogeneity because it includes specific dimensions. Researchers working from a HiTOP perspective are able to focus on homogeneous dimensions at lower levels of the hierarchy if they want to pinpoint the specific processes that connect psychopathology and outcomes of interest. This degree of resolution is impossible with almost all *DSM* and *ICD* diagnoses, which combine constellations of loosely related symptoms and traits into individual diagnostic constructs (Smith, McCarthy, & Zapolski, 2009).

Worked Example

This tutorial uses the Inventory of Depression and Anxiety Symptoms (IDAS), mentioned earlier, to examine the fundamental components of the internalizing domain as they relate to distress tolerance. Distress tolerance can be defined as the ability to maintain goal-directed activity in the face of discomfort, such as pain, traumatic memories, or obsessive thoughts (Leyro, Zvolensky, & Bernstein, 2010). Like other stress-reactivity outcomes, distress tolerance is empirically related to an array of internalizing conditions, including major depression, generalized anxiety disorder, and social phobia (e.g., Macatee, Capron, Guthrie, Schmidt, & Cougle, 2015). We analyzed the data without *a priori* hypotheses regarding the relative effect sizes of specific symptom dimensions, although we generally expected that greater severity of internalizing symptoms would be correlated with poorer distress tolerance.

Measures. In this example, we used a dataset of 145 university students who completed the IDAS and a self-report measure of distress tolerance, described below. The IDAS assesses 10 specific symptom dimensions from the internalizing domain: suicidality, lassitude, insomnia, appetite loss, appetite gain, ill temper, social anxiety, panic, traumatic intrusions, and well-being (cf. the bottom tier of Figure 1). It also features 2 broad scales that tap into the general features

of internalizing problems: general depression and dysphoria. For simplicity, we omitted general depression from our analysis because it is highly correlated (r = .95) with dysphoria in this sample, and, unlike dysphoria, contains items that overlap with the 10 lower-order scales. Prior research has confirmed that the IDAS has favorable psychometric properties in diverse samples (e.g., Nelson, O'Hara, & Watson, 2018). The 10 specific scales are reasonably, but not excessively, intercorrelated, and in factor analyses they define a single underlying dimension (Watson et al., 2007).

Distress tolerance was assessed using the Distress Tolerance Scale (Simons & Gaher, 2005). This 15-item self-report inventory provides an overall index of how effectively people feel that they manage upsetting emotional states.

Results. Table S1 presents descriptive statistics. The sample means for all IDAS scales closely parallel those previously reported in university-student samples (Stasik-O'Brien et al., 2019). Correlations among the 10 lower-order scales range from .02 to .67 (median r = .33) (after reversing the sign of correlations between well-being and other scales).

To examine the effects for each IDAS dimension in isolation, we performed a series of simple (i.e., single-predictor) regressions of distress tolerance on IDAS scales. Table 1 shows that the standardized effects ranged from -.60 to .39, and all were statistically significant at the .05 alpha level. The effect for dysphoria was largest by a substantial margin (standardized effect = -.60), whereas ill-temper, insomnia, and appetite gain (-.20 to -.27) had the most modest associations with self-reported distress tolerance. We next performed a multiple regression, which included all 10 lower-order scales as predictors—but omitted Dysphoria to avoid collinearity problems—to examine each effect holding the others constant. Most effect sizes dropped appreciably in this analysis, consistent with the idea that all subscales are indicators of

the same broader construct. None exceeded |.20|, and all were smaller than |.15| except panic and well-being, an index of positive affectivity at one pole and anhedonia at the other.[2] This general attenuation of effects suggests that there is redundancy among these components of internalizing problems, and that most of their external validity may be accounted for by the core disposition (internalizing) that they share.

Interpretation. We conclude that many specific aspects of the internalizing domain are meaningfully related to self-reported distress tolerance. The largest effect was observed for the dysphoria scale, which was empirically derived to reflect the higher-order internalizing dimension (Watson et al., 2007). Effects for the lower-order symptom components were smaller, and there was some variability among them. Traumatic intrusions, lassitude, and panic symptoms exhibited the largest correlations with distress tolerance, whereas ill-temper, insomnia, and appetite gain had the smallest. There was some evidence from the multiple regression that well-being had an independent association with the outcome, whereas other dimensions had very small unique effects.

For published projects that follow this general approach, see Waszczuk, Zimmerman et al. (2017), Watson, Clark, Chmielewski, & Kotov (2013), and Watson et al. (2007).

Tutorial 2: Criterion Validity of Higher-order Dimensions

We have emphasized that the HiTOP model organizes clinical phenotypes hierarchically. In this rubric, no single level optimally represents psychopathology. Broader dimensions are

[2] We caution researchers who perform this type of analysis about possible interpretative problems when including many predictors in the same analysis. Statistical adjustment for correlated predictors in a multiple regression context can sometimes change the meaning or content of an independent variable, especially when predictor correlations are large (Lynam, Hoyle, & Newman, 2006). In other words, the "partialled" variable could be a poor reflection of the target construct. Researchers can examine collinearity diagnostics to get a sense of how problematic predictor intercorrelations are. In our multiple regression, the variance inflation factor ranged from 1.16 to 2.48 for the 10 IDAS scales, suggesting that there was not prohibitive collinearity among the predictor scales (see Pedhazur, 1997, chapter 10).

bound to be ideal for some research contexts, and narrower dimensions for others; this tension is akin to the bandwidth-fidelity dilemma in personality psychology (e.g., Cronbach & Gleser, 1957). We expect that, without the benefit of strong theory or prior evidence, researchers will find it difficult to intuit the level(s) of analysis that will have the largest empirical association with some risk factor or outcome. Therefore, we argue that researchers ought to test empirically how external correlates (e.g., etiological markers) map onto psychopathology at various levels of the hierarchy.

Suppose an investigator theorizes that depression predicts loafing in the workplace. A conventional way to test this idea in observational research is to estimate the correlation between loafing and major depressive disorder symptoms. We know, however, that syndromes like major depression reflect a mixture of both broad (e.g., internalizing, distress) and specific (e.g., anhedonia, hostility) individual difference factors. Any of these might be most proximally linked to loafing. The sizes of these influences are impossible to know (and compare) until the researcher examines them directly.

Further, the literature on these types of research questions often develops in disorder-specific silos. So whereas one investigator may have found an association between loafing and major depression symptoms, another may have found an association between loafing and GAD (or panic, social anxiety, and so on). Understanding the extent to which these associations are specific to one construct over another, in comparison to the extent to which they can be parsimoniously accounted for by higher-order dimensions, can have important implications for developing appropriately targeted interventions (Forbes, Rapee, & Krueger, 2019).

In this tutorial, we examine the emotional correlates of healthy romantic relationships. We first establish that a hierarchical, dimensional model makes sense of the pattern of

covariation among anxiety and depression symptom scales. We then test associations between relationship functioning and dimensional phenotypes across three levels of a hierarchical model of anxiety and depression, differentiating the specificity versus generality of each association. This approach allows us to evaluate the extent to which each construct in the hierarchical model is associated with the outcome of interest (i.e., a total effect), and then to disentangle the portion of that relationship that is unique to the construct of interest (i.e., a direct effect) versus accounted for by higher-order factors that capture variance shared with other constructs in the model (i.e., an indirect effect). This method can thus transcend multiple disorder-specific literatures, and substantially advance our understanding of the hierarchical structure of the risks and outcomes of psychopathology.

Worked Example

People in satisfying romantic relationships report fewer emotional problems. Research across developmental, social/personality, and clinical fields confirms that the quality of romantic relationships is inversely associated with dysphoria, panic, worry, social anxiety, and the primary features of several other emotional disorders (e.g., South, in press). This non-specificity could reflect an association of relationship quality with broad psychopathology dimensions—such as fear, distress, internalizing, and even the general factor of psychopathology—that are manifested in various diagnosable emotional problems (e.g., South & Krueger, 2008). On the other hand, it could reflect multiple independent associations of relationship quality with specific emotional disorder symptoms or syndromes (e.g., panic, social anxiety).

In this example, we examine relationship quality in association with HiTOP syndrome-, subfactor-, and spectrum-level constructs. We determine the degree to which relationship quality's observed associations with the various syndromes are explained (i.e., statistically

mediated) by higher-order dimensions. The dataset for this tutorial comes from a larger survey of romantic experiences and mental health in an Australian community sample (see Wong & Forbes, 2020). We selected a subset of respondents (N = 725) who reported that they currently were in a romantic relationship.

Measures. Emotional problems were assessed using popular surveys for depression, generalized anxiety disorder, social phobia, panic disorder, and obsessive-compulsive disorder. These were the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer & Williams, 2001), Brief Measure for Assessing Generalized Anxiety Disorder (GAD-7; Spitzer, Kronenke, Williams & Lowe, 2006), Social Phobia Inventory (SPIN; Connor et al., 2000), Panic Disorder Severity Scale Self-Report (PDSS-SR; Newman, Holmes, Zuellig, Kachlin, & Behar, 2006), and Obsessive-Compulsive Inventory-Revised (OCI-R; Foa et al., 2002), respectively. All measures provided a continuous index of symptom severity. Relationship quality was assessed using the 6-item Modified Quality of Marriage Index (QMI; Nazarinia, Schumm & White, 2009). Despite its name, the QMI is often used to assess the quality of *non-marital* romantic relationships, as we do in this example. Respondents rated their agreement with a series of statements about relationship satisfaction over the past 2 weeks (e.g., *I really feel like part of a team with my partner*) on a 6-point Likert scale. We used the QMI total score (i.e., mean of all item responses) as our index of relationship quality in this analysis.

Results. The first part of the analysis involved creating latent dimensions from the set of observed symptom measures using CFA. In factor-analytic language, this step generates the *measurement model*. Following the structure represented in Figure 1, we hypothesized that a fear dimension would account for the correlations among social phobia, panic, and obsessive-compulsive symptoms, whereas a distress dimension would account for the correlation between

depression and generalized anxiety. Table S3 reports the scale intercorrelations. The fear and distress dimensions are called *first-order* factors because they are just one step removed from the observed variables. Figure 1 also led us to expect that the fear and distress factors would be substantially correlated, reflecting a shared superordinate internalizing factor, which here represents a *second-order* factor.[3]

This hypothesized model fit the data well. Indices of model fit suggested that the internalizing, distress, and fear factors adequately represented the correlations among symptom scales: $\chi^2(4) = 9.23$, $p = .06$; CFI = 0.99; RMSEA = 0.042; SRMR = 0.015 (see Brown, 2015, chapter 5, for details on how to evaluate factor models). The standardized factor loadings on the first-order factors ranged from 0.73 to 0.88, and the loadings of fear and distress on the internalizing factor were both 0.93. It was necessary to constrain the two (unstandardized) second-order loadings to be equivalent for model identification purposes (see Brown, 2015, chapter 3, for information on identification).

The second step in the analysis involved regressing psychopathology variables on relationship quality. To answer our question about potential etiological pathways across multiple levels of the dimensional hierarchy, we estimated total, direct, and indirect effects of relationship quality on psychopathology outcomes (i.e., symptoms of depression, GAD, panic, social anxiety, and OCD; distress and fear; and internalizing) one at a time. Figure 2 diagrams the distinction between direct and indirect effects in the context of our factor model for this dataset. The *total*

[3] Another way to parameterize a higher-order internalizing factor here would have been to include use a bifactor model (e.g., Bornovalova, Choate, Fatimah, Petersen, & Wiernik, 2020; Lahey, Krueger, Rathouz, Waldman, & Zald, 2017). In bifactor (also called hierarchical) models, all indicators load onto a single general factor and only one of a set of group factors. Broadly speaking, the general factor represents the commonality among all observed variables, whereas group factors represent more specific sources of variation among these variables after they are residualized on the general factor. Bifactor models are increasingly popular in structural research on psychopathology, although there are signs that they have been applied uncritically or misinterpreted in this literature (e.g., Watts, Poore, & Waldman, 2019).

effect represents the bivariate (or "zero-order") association between relationship quality and a psychopathology dimension. This effect does not adjust for shared variance between the psychopathology variable (e.g., panic symptoms) and the other variables in the model. Next, the *direct effect* represents the extent to which relationship quality and a psychopathology dimension are related, above and beyond any higher-order constructs (e.g., fear and internalizing). Thus, it gives the proportion of the total effect that is not statistically mediated by higher-order psychopathology dimensions. For example, the direct effect of relationship quality on depression occurs independently of the internalizing and distress dimensions; it reflects the association between relationship quality and the part of depression that is independent of other emotional problems.

In contrast, the *indirect effect* reflects the proportion of the association between relationship quality and psychopathology that *is* mediated by higher-order constructs.[4] Note that there is no indirect effect of relationship quality on internalizing in this example, because internalizing is at the apex of the hierarchy in our model. For more conceptual and quantitative information on total, direct, and indirect effects in path analysis (the foundation of factor analysis and structural equation modeling), we recommend textbooks on applied multiple regression/correlation analysis (e.g., Pedhazur, 1997, chapter 18).[5]

The first 5 rows of Table 2 present the total effects of relationship quality on the symptom measures. All coefficients were negative (standardized effect range = -.24 to -.37) and

[4] The standard errors of the indirect effects have an asymmetric distribution (Cheung & Lau, 2008), so bias-corrected bootstrap confidence intervals were estimated using 500 bootstrap samples and a maximum likelihood estimator. The resulting intervals were nearly identical to those based on maximum likelihood estimation with robust standard errors.

[5] We note that we are working with cross-sectional data, so tests of mediation require strong assumptions about the direction of causality. A model that assumes a reverse causal sequence (i.e., psychopathology predicts relationship quality) would fit the data equally well (see MacCallum, Wegener, Uchino, & Fabrigar, 1993).

indicated that, as expected, healthy relationships were linked with fewer emotional problems. The total effects for fear, distress, and internalizing were -.32, -.36, and -.37, respectively. Again, we can conceptualize these effects as the bivariate correlations between relationship quality and psychopathology constructs.

Table 2 also lists the direct effects of relationship quality on the observed symptom outcomes. As noted above, the dependent variable in these regressions is the part of symptom scales that is *not shared with other outcomes*. This unshared, or unique, variation often is called the "residual variance" in factor analysis and structural equation modeling more generally. Likewise, the regressions of fear and distress on relationship quality reflect associations that are adjusted for the overlap each first-order factor has with the second-order internalizing factor. That is, the outcomes in these first-order factor regressions represent pathology that fear does not share with distress (and vice versa).

All direct effects for syndrome- and subfactor-level constructs were small, compared to the total effects. Three out of five syndrome-level direct effects were approximately 0. However, the (negative) direct effect on depression was statistically significant. This result indicates that relationship health was (negatively) associated with the part of depression that remained after adjusting for the overlap depression has with other emotional problems. Stated differently, there is some, albeit modest, connection between relationship quality and depression *even after* taking into account what is shared between depression and internalizing more broadly. There was also a statistically significant direct effect on generalized anxiety symptoms. This effect was *positive,* suggesting that relationship quality was positively associated with the part of generalized anxiety that remained after adjusting for the higher-order dimensions. Notably, this positive direct effect with generalized anxiety symptoms is a suppressor effect (i.e., the partialled

effect reversed in sign compared to the zero-order effect; see Watson, Clark, Chmielewski, & Kotov, 2013); this positive direct effect is a corollary of the negative direct effect found for depression, as there were only two indicators for the distress latent variable here.[6]

The final column in Table 2 presents the proportion of the total effect for each outcome that was accounted for by factors higher in the hierarchy. For symptom scales, this column indicates the total effect accounted for by distress, fear, and internalizing factors. For distress and fear factors, this column indicates the total effect accounted for by internalizing. This proportion, in turn, represents the magnitude of the indirect effect, and it is computed by subtracting the direct effect from the total effect and dividing the difference by the total effect. These values tell us that for all symptom scales—with the exception of depression—the indirect effect (of relationship quality via the fear, distress, and internalizing factors) explained the large majority of the bivariate association with relationship quality. This same pattern was evident for the first-order factors. Regarding depression, roughly one-third of the total effect was *not* mediated by the factors (i.e., was unique to depression).

Interpretation. According to our analysis, much of the association between relationship quality and emotional problems operated at a higher-order level. Put another way, after adjusting for higher-order dimensions, there were relatively small and inconsistent associations between relationship quality and the *DSM* syndromes that clinicians tend to diagnose in routine practice.

[6] For demonstration purposes, we kept this factor model simple. In practice, it is useful to have more than 2 indicators per factor. In our case, for instance, the fact that we had only 2 (latent) indicators of the internalizing factor—whose loadings were constrained to equality (for model identification)—meant that the direct effects of relationship quality on these 2 first-order factors were bound to be equal in size (and opposite in sign). We saw a similar instance of countervailing direct effect estimates for the 2 indicators of the distress dimension (depression and generalized anxiety). Greater representation of possible factor content (i.e., observable manifestations of the factors; in our case, symptom dimensions in each psychopathology domain)—for both first- and higher-order factors—would avoid this type of artifact.

This pattern of effects highlights a key methodological issue. The modal study of the correlation between close relationships and mental health involves a single disorder entity (see Egan & Smith, in press), such as social anxiety. But here we see that any associations of relationship functioning with any specific manifestation of the internalizing dimension (e.g., social anxiety symptoms) might reflect a broader association between relationship functioning and higher-order dimensions such as internalizing (e.g., South, Krueger, & Iacono, 2011). Reframing studies around a hierarchical model of psychopathology can therefore give a more complete account of how potential causal factors relate to mental health problems. As a result, theory becomes more nuanced. Based on the present results, for example, we might posit that satisfying relationships buffer against all internalizing problems and, at a finer level of resolution, against depression-specific problems too. This is more complex (and accurate) than hypotheses guided by *DSM* alone (e.g., "satisfying relationships mitigate risk for major depression"). Moreover, this theoretical development feeds into improved research design. That is, study design is more efficient to the extent that, based on prior results, researchers are able to target the emotional phenotypes most relevant to relationship functioning.

Although broad dimensions, such as internalizing, seem to have strong criterion validity across most published investigations, research on hierarchical models is bound to discover important roles of lower-order dimensions. As we alluded to, depressive symptoms in our analysis remained significantly correlated with relationship quality after adjusting for the pathology depression shares with other emotional problems. (Note, however, that depression's higher-order (i.e., indirect) effect was twice as large as the depression-specific (i.e., direct) effect.) This result supports the idea that multiple levels of breadth can be important to consider in mental health research, depending on the association of interest. For instance, future research

might benefit from testing both higher-order *and* depression-specific pathways between relationship quality and depression.

For published examples of this multi-tiered approach to criterion-validity testing, see Forbes, Magson, & Rapee (2020), Hamlat, Young, and Hankin (2019), Rodriguez-Seijas, Stohl, Hasin, and Eaton (2015), Sellbom, Carragher, Sunderland, Calear, and Batterham (2020). Also, in this article's supplemental material, we recast Tutorial 1 as a latent variable model and demonstrate one way to estimate the criterion-validity effects of lower-order symptom dimensions over and above a common factor representing the broad internalizing dimension.

Tutorial 3: Developmental Trends in Latent Psychopathology Dimensions in Adolescence

Developmentalists have been studying internalizing and externalizing variation for decades, dating back to Achenbach's early factor analyses of child mental disorder symptoms (Achenbach, 1966). This body of research thus provides a guide for dimensional research across the entire life span (Beauchaine & McNulty, 2013). However, much of the evidence supporting the HiTOP model comes from cross-sectional studies of adults. Less is known about how HiTOP dimensions change over time, particularly in developmental stages associated with rapid shifts in risk for psychopathology such as adolescence. Longitudinal research is needed to test the extent to which dimensional phenomena are developmentally coherent to ensure that developmental changes in the structure of psychopathology are captured in the HiTOP model (Caspi et al., 2020; Forbes et al., 2016).

Recent work has used longitudinal structural equation modeling techniques to identify continuity and change in psychopathology phenotypes—especially internalizing, externalizing, and the *p*-factor—in children (Olino et al., 2018), adolescents (Laceulle, Vollebergh, & Ormel, 2015; Snyder, Young, & Hankin, 2017), and older adults (Eaton, Krueger, & Oltmanns, 2011).

Formally testing longitudinal measurement invariance is important because it determines whether observed changes in latent dimensions can be interpreted as genuine change in that dimension over time, or reflects fluctuations in measurement properties of the observed variables (Horn & McArdle, 1992). Once such measurement invariance is established, longitudinal research can further characterize the factors that predict the natural course of psychopathology and its downstream consequences.

Worked Example

The transition to adolescence marks a period of social upheaval. Youth tend to distance themselves from parents and invest heavily in peer relationships as a source of social and emotional support (Steinberg & Morris, 2001). Fostering friendships during adolescence is therefore a key developmental milestone.

Anxiety and depression can derail this process. Cross-sectional data tell us that youth diagnosed with emotional disorders report less satisfying peer relationships and more difficulty establishing and keeping friendships (Rapee et al., 2019). Sociometric studies document that anxious and depressed children and adolescents tend to be at the periphery of friendship groups (Prinstein & Giletta, 2016). These interpersonal processes affect, and are affected by, social competence—self-perceptions of one's ability to have meaningful social relationships (Harter, 1999).

These observations may reflect a dynamic process connecting emotional problems and self-perceptions of social functioning. That is, anxiety and depression may limit social prospects, and peer relationship difficulties may in turn exacerbate feelings of anxiety and depression (La Greca & Harrison, 2005). This type of positive feedback loop portends an intractable course of anxiety and depression in adolescence and beyond.

We put this developmental psychopathological hypothesis to the test by examining dynamics between internalizing problems and social competence over three occasions, each separated by about 18 months, in a community sample at the transition to adolescence. We defined internalizing as the commonality among adolescents' self-reports on three frequently used assessment instruments for anxiety and depressive symptoms. We hypothesized that self-perceived social competence and emotional problems would reciprocally influence one another over time, consistent with the idea of a positive feedback loop. The data for this tutorial come from the Genes, Environment, and Mood Study (see Hankin et al., 2015, for details). Participants (N = 682; 381 girls) were on average 12 years old (SD = 2.38) at wave 1. About two-thirds were white, 11% black, 9% Asian or Pacific Islander, 1% American Indian or Alaska Native, and 11% identified as multi-racial or another race. Thirteen percent identified as Latinx.

As a precursor to testing these prospective effects, we investigated longitudinal measurement invariance in our internalizing construct, which was represented by three observed indicators of anxiety and depressive symptoms.

Measures. Anxiety and depression were assessed with popular instruments that are well-validated with regard to inferences about children's internalizing symptoms: self-report of Children's Depression Inventory (CDI; Kovacs, 1985), self-report of Manifest Anxiety Scale for Children (MASC; March, Parker, Sullivan, Stallings, & Conners, 1997), and caregiver report of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001). Our analysis involved total scale scores from the CDI and MASC, and the SDQ Emotional Symptoms subscale score, which reflects a mixture of anxiety and depressive symptoms (e.g., "I worry a lot"; "I am often unhappy, depressed, or tearful"). Higher scores on all three measures reflect more severe symptoms.

The Self-Perception Profile for Adolescents (Harter, 2012) was administered to measure youth attitudes about themselves along a number of dimensions, ranging from athletic ability to physical appearance to school performance. In the current analysis, we focused on the social competence subscale, which asks respondents to rate the extent to which they match one end or another of six bipolar items tapping into the number and quality of their friendships with peers (e.g., "some kids find it hard to make friends" versus "other kids find it's pretty easy to make friends"). Higher social competence scores reflect better functioning.

Results. Table S5 presents descriptive statistics for our observed variables. All contemporaneous indices of internalizing problems were significantly correlated with one another (*r*s = .15 to .43), consistent with our plan to treat them as expressions of the same underlying factor. Social competence was also modestly-to-moderately related to observed internalizing symptoms cross-sectionally (*r*s = -.15 to -.48) and at adjacent waves (*r*s = -.12 to -.33), confirming the link between these two constructs.

We used latent variable modeling to define an internalizing factor as the common variance among CDI, MASC, and SDQ scores at all three assessment waves. We applied confirmatory factor analysis (CFA) to test for measurement invariance, a precondition for the cross-lagged panel models (CLPMs) we estimated to test our primary hypotheses (Horn & McArdle, 1992), of this longitudinal model. In the first step in this process, we tested a longitudinal CFA, in which all factor loadings were freely estimated, called the unconstrained model. This model was a good fit to the data, $\chi^2(15) = 33.23$, $p < .01$; CFI = 0.99; RMSEA = 0.046; SRMR = 0.048, suggesting that the assumption of configural invariance—the same three observed variables serve as indicators of the internalizing factors across waves—was satisfied. All loadings were statistically significant and substantial, except for SDQ, which was slightly

below conventional thresholds at the first two waves (standardized factor loadings at waves 1-3 = .28, .27, .38).

Next, in a test of loading invariance, we fit a new model that constrained loadings of corresponding indicators (e.g., CDI_{wave1}, CDI_{wave2}, CDI_{wave3}) to take on the same (unstandardized) values across measurement occasions. This restriction did not significantly degrade model fit, $\chi^2_{diff}(4) = 2.28$, $p = .68$, indicating that the three observed indicators measured the latent construct in approximately the same way over time. That is, the association between the internalizing factor and indicators did not change appreciably across waves. After confirming metric invariance, we could move forward with tests of how internalizing and social competence shape one another over time.

There are many ways to study the temporal dynamics of two related constructs (Usami, Murayama, & Hamaker, 2019). The key is to choose the analysis plan that most closely maps onto the developmental question at hand. In our case, we aimed to study how individual differences in internalizing problems prospectively predicted changes in self-perceived social competence (and vice versa). This question was a close match for the CLPM, which is a popular choice for developmental researchers studying co-development among two related processes (Orth, Clark, Donnellan, & Robins, 2020). The CLPM requires that two constructs are measured on at least two occasions, and it estimates the prospective effects of each construct on the other, above and beyond the observed within-construct continuity in each process (McArdle, 2009).

The CLPM fit the data well, $\chi^2(41) = 68.21$, $p < .01$; CFI = 0.99; RMSEA = 0.034; SRMR = 0.047. It included both autoregressive (e.g., wave 2 social competence regressed on wave 1 social competence) and cross-lagged (e.g., wave 2 social competence regressed on wave

1 internalizing) paths, which represent continuity and reciprocal influences, respectively (see Figure 3). The complete list of parameter estimates is presented in Table S6.

To foster model parsimony, we evaluated several sets of equality constraints that constrained parameters to the same value across waves, similar to the restrictions imposed in the test of loading invariance (see Curran & Bollen, 2001). This process resulted in fixing autoregressive, cross-lagged, and residual covariance paths to equality over time, except the prospective effect of social competence on internalizing, which was freely estimated (see supplementary material for details).

This final model, represented in Figure 3, also fit the data well, $\chi^2(48) = 76.52$, $p < .01$; CFI = 0.99; RMSEA = 0.032, SRMR = .049. The autoregressive paths for internalizing were fairly large (standardized effects = .69 to .74), whereas they were more moderate for social competence (.42 to .44). There were effects of internalizing variation on subsequent social competence (-.23 to -.25) above and beyond the temporal stability of both constructs. The reverse effects (i.e., social competence on internalizing), represented in Figure 3 by dashed lines, were small and not statistically significantly different from zero (-.01 to .13).

Interpretation. There were two main findings here. First, through measurement invariance tests, we observed that there was continuity in how the latent internalizing construct was expressed across three years in adolescence. Not only did a three-indicator measurement model adequately reproduce the correlations among observed anxiety and depression variables across waves (configural invariance), but the different indicators related to the latent construct equivalently over time (metric invariance). This result suggests that internalizing, as a latent construct, was developmentally coherent through early adolescence when operationalized via these self-report symptom inventories.

Further, levels of internalizing problems in this age group showed substantial rank-order stability in individual differences over the three-year span, as seen in the autoregression coefficients (standardized effects of approximately .70). Other research has reported similar estimates over intervals this long (and longer) (Krueger et al., 1998; Sunderland, Slade, Carragher, Batterham, & Buchan, 2013; Vollebergh et al., 2001). Much more investigation is needed into normative patterns of stability and change of HiTOP dimensions, and how they align with results drawn from research on the time course of categorical diagnoses.

Second, the CLPM revealed that there is overlap in the development of internalizing and social competence. The two were strongly related cross-sectionally in this sample, as evidenced, for example, by a model-implied correlation of -.57 between the two at wave 1. Moreover, internalizing prospectively predicted decrements in social competence even after adjusting for continuity in both constructs across waves. These results tie into a robust literature on the negative effects of anxiety and depression on self-perceptions of ability to engage in and have meaningful peer relationships in adolescence (Harter, 1999). The reverse was not true: social competence did not predict changes in internalizing levels. This null effect could signal that social competence is not an important predictor of internalizing trajectories, but it might also be a function marked temporal stability in the internalizing spectrum mentioned above, given there is comparatively little variation left to predict over time.

This analytic framework can be applied to understanding the developmental coherence and longitudinal stability of HiTOP dimensions—and more broadly to characterizing the risks, consequences, and reciprocal relationships among HiTOP dimensions and their correlates across the lifespan.

Common Questions and Challenges

Is HiTOP For Self-report Measures Only?

HiTOP grew out of factor analysis of self-reported and interviewer-rated symptoms and diagnoses. These were the data available in large datasets that were well suited to the type of latent variable modeling traditionally used to resolve the correlations among latent dimensions. For example, investigators have repeatedly mined the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), a multiwave study involving interviewer ratings of mental disorder symptoms in over 40,000 US adults, for information on the empirical structure of psychopathology.

This need not be the predominant strategy for research on the utility of HiTOP constructs. First, informant reports are known to have incremental validity and, for some outcomes, superior predictive utility, relative to a proband's own account of mental health problems. Indeed, Achenbach's seminal work on dimensional models of child psychopathology relied in large part on collateral sources such as parents, caregivers, and teachers (reviewed in Achenbach, 2020). This trend continues in current HiTOP-informed research (see, e.g., Boudreaux, South, & Oltmanns, 2019).

Second, we recommend greater integration of psychopathology dimensions with experimental and laboratory-based designs involving behavioral observation, task performance, psychophysiology, and neuroimaging (e.g., Latzman, DeYoung, & the HiTOP Neurobiological Foundations Workgroup, in press; Snyder, Friedman, & Hankin, 2019). This type of research can determine whether HiTOP dimensions relate more robustly than categorical diagnoses to behavioral and biological vulnerabilities.

For instance, there is a robust line of clinical research on fear conditioning in anxious populations. Anxious people exhibit more intense defensive responses to safety cues during

conditioning (Duits et al., 2015). But it remains unknown whether disrupted fear learning is a general characteristic of all internalizing problems, correlated predominantly with fear-based disorders (e.g., phobias), or tied to specific diagnoses (e.g., PTSD) (e.g., Craske et al., 2009). A hierarchical, dimensional approach to psychopathology assessment—as reflected in, for example, the IDAS (see Tutorial 2)—could resolve this issue, significantly advancing etiological theories of anxiety-based disorders.

Third, the concordance of verbal-report and other measures of psychopathology can be examined by integrating multiple measurement modalities into the same structural model. In one recent study, Hill, South, Egan, and Foti (2019) modeled a latent dimension of "neural reactivity" using event-related potentials derived from different experimental conditions and examined the correlations between this factor and a latent internalizing factor. In this same vein, Yancey et al. (2016) modeled the shared variance of heart rate, startle potentiation, and corrugator muscle tension expressed during an affective picture-viewing task along with verbal reports of fear and fearlessness. They found that this latent factor was meaningfully correlated with symptoms of DSM-defined fear disorders and other physiological risk factors.

Does HiTOP Require Large Samples?

The short answer is "not necessarily." As mentioned, much of the work on the latent structure of psychopathology has relied on epidemiological studies, partly because they tend to assess broad swaths of psychopathology. Using HiTOP constructs to test hypotheses in other fields is a different undertaking, and sample size requirements can vary widely depending on the research context.

The sample size necessary to estimate factor analytic and structural equation models depends on the number of observed variables, the number of latent variables, the proportion of

missing data, and the magnitude of both factor loadings and paths between latent variables (Wolf, Harrington, Clark, & Miller, 2013). Simple models with clearly defined factors and strong loadings can be estimated with sample sizes between 50 and 200 (Boomsma, 1983; MacCallum et al., 1999). In the least favorable conditions, these models may require sample sizes up to 450 and beyond. Should latent variable models not converge, or should the investigator simply prefer a simpler method, it is possible to test the same hypotheses without using latent variable models. This is accomplished by estimating a path model between scale scores from tests that assess HiTOP constructs directly. This is the approach taken in Tutorial 1, for example. A list of HiTOP-friendly measures is available at https://hitop.unt.edu/clinical-tools/hitop-friendly-measures. Even in small samples, however, it is worth attempting to estimate latent variable models. This is because path models assume that measures are perfectly reliable—an assumption that is almost never true—which attenuates power and increases the chance of errors (Cole & Preacher, 2014). When measures are unreliable, a structural equation model can detect effects half as large as those that could be detected using a path model (Wolf et al., 2013). Further, investigators might have hypotheses about an effect that is moderate or large; power analysis in SEM will often focus on power for a specific parameter, not the overall fit of the model.

Researchers planning to use significance tests should perform power analyses for all study designs, including those involving HiTOP. When only observed variables are involved, the familiar power analysis methods will suffice. The `pwr` package in R (Champely, 2020) implements many of the recommendations in Cohen (2013). For instance, in Tutorial 1, we would determine the sample size needed to detect the change in variance explained when adding

the 10 lower -order IDAS dimensions to a multiple regression that also included the internalizing dimension (assuming this was our primary hypothesis).

For investigators who want to compute power for effects in a structural equation model, there is a growing number of user-friendly tutorials and dedicated software packages that provide guidance (Wolf et al., 2013). Within the R software language, there are several recently developed (and evolving) packages dedicated to power computations (see, e.g., Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2020). The Mplus software has similar capabilities (e.g., Muthén & Muthén, 2002), as does SAS (MacCallum, Browne, & Sugawara, 1996).

Does HiTOP Work Only In Certain Populations?

There is no ideal population for HiTOP-informed research. Again, a fair amount of work thus far has been conducted in epidemiological datasets, where large samples afford much more precision around parameter estimates. Yet basic architectural elements of the HiTOP model appear to be replicable across more specific populations, including patient groups (reviewed in Kotov et al., 2017). Much of the structure also seems to be invariant across gender, developmental stage, and region of the world (Krueger et al., 2018).

We also note that one of HiTOP's central tenets is that psychopathology phenotypes are dimensional. Thus, every person theoretically can be characterized by their location on a spectrum representing, say, mood lability, fear, or disinhibited externalizing. Very fine gradients are possible, such that meaningful individual differences can be detected in most, if not all, populations.

Is HiTOP Compatible With The Research Domain Criteria?

Beginning in 2007, the US National Institute of Mental Health developed the Research Domain Criteria (RDoC) initiative as an alternative research framework to *DSM*. RDoC is not a

classification system per se, but a research heuristic that identifies biological and behavioral processes relevant to mental disorders (Kozak & Cuthbert, 2016). The RDoC framework is operationalized in a matrix (https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/constructs/rdoc-matrix.shtml) that crosses eight units of analysis (genes, molecules, cells, circuits, physiology, behavior, paradigms, self-report) with six constructs (negative valence, positive valence, cognitive, social, arousal and regulatory, and sensorimotor systems).

There are some salient conceptual differences between HiTOP and RDoC (see Clark et al., 2017, for an extended comparison of RDoC with current diagnostic systems, which overlaps with a HiTOP-RDoC comparison in several ways). HiTOP is oriented around the subjective signs and symptoms of mental disorder that drive patients to get treatment and that practitioners encounter regularly in the clinic, whereas RDoC, although it does not exclude such data, emphasizes covert aspects of mental disorder risk that are not key pieces of disorder phenomenology or psychological theories. For instance, RDoC characterizes mental health problems as "disorders of neural circuits" (Insel & Cuthbert, 2015). Much of the emphasis appears to be on the neurobiological level of analysis.

Despite their different emphases, there is emerging consensus that these two systems can work in concert for a more comprehensive understanding of psychopathology. HiTOP provides a bridge from biobehavioral processes codified in RDoC to clinically recognizable phenotypes. This link makes RDoC, which was not designed with clinical implementation as a near-term goal, much more practically relevant. At the same time, the research paradigms and basic science discoveries motivated by RDoC can identify the biological building blocks of the clinical constructs represented in the HiTOP model. With these goals in mind, recent scholarship from the HiTOP consortium has focused on developing a detailed interface between HiTOP and

RDoC (Michelini, Palumbo, DeYoung, Latzman, & Kotov, 2020; Perkins, Joyner et al., 2020; Perkins, Latzman, & Patrick, 2020).

How Does HiTOP Address Development?

Much of the literature underpinning the current HiTOP model has been conducted in cross-sectional studies and adult samples, focusing on an incomplete range of psychopathology domains (Kotov et al., 2017). This has resulted in two substantial limitations when applying the HiTOP framework in developmental psychopathology research. First, the current framework has a static structure that does not account for potential changes in the empirical structure of psychopathology across development. For example, predispositions towards general psychopathology are evident in early childhood temperament, with propensities towards internalizing versus externalizing domains of psychopathology emerging later in childhood, and narrower and more varied symptoms and syndromes crystalizing in adolescence and early adulthood (see Forbes, Rapee, & Krueger, 2019, for a review). This may well manifest as non-invariance in dimensional models of psychopathology throughout development (cf. the varying structure of the ASEBA inventories by age group; Achenbach, 2020). Studies that have examined the structure of psychopathology longitudinally—and across multiple developmental periods—have tended to converge on dimensions represented in the HiTOP model (Caspi et al., 2020; see Forbes et al., 2016, for a review). However, the extent to which these dimensions demonstrate longitudinal measurement invariance throughout development requires more systematic and comprehensive efforts, and is a core focus of the HiTOP Developmental Workgroup (Kotov et al., under review).

Second, the current working model spans approximately two-thirds of the diagnoses in the *DSM-5*, but does not yet include important domains of child-onset psychopathology, such as

autism spectrum disorder and other neurodevelopmental disorders. This limits the scope of research that can currently be accommodated in a HiTOP framework. The HiTOP Revisions Workgroup is currently working on ensuring that the model remains up-to-date with current research, and the first formal revision of the model is anticipated to incorporate a broader variety of psychopathology—including autism-related symptoms (cf. Michelini et al., 2019; Noordhof et al., 2015).

Conclusions

Investigators across diverse fields and training backgrounds seem to be wondering whether categorical diagnoses have outlived their utility for most mental health research. Is HiTOP a viable alternative? We believe so, but we acknowledge that this is ultimately an empirical question that must be tested scientifically. This article presented tutorials intended to help investigators pilot this framework—and evaluate its utility—in various applied research contexts.

We recognize that research habits are hard to break. Funding bodies, professional organizations, and training programs for many years institutionalized categorical disorders as research targets, notwithstanding some recent trends in the opposite direction (e.g., RDoC). We contend, however, that the practical benefits of a hierarchical, dimensional approach are worth the effort. HiTOP promises to unify unnecessarily fragmented research literatures, improve precision in psychological theories, and generate new substantive insights into the causes and consequences of mental illness.

The HiTOP framework is relevant to all researchers who study mental illness, not just those who specialize in diagnosis and assessment. We hope this article will help make HiTOP more accessible and promote widespread empirical evaluation of HiTOP's utility as a research

tool, with the potential of improving clinical practices and, in turn, clinical outcomes for people with mental health problems.

References

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs, 80,* 1-37.

Achenbach, T. M. (2009). *The Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications*. University of Vermont, Research Center for Children, Youth, & Families.

Achenbach, T. M. (2020). Bottom-up and top-down paradigms for psychopathology: A half-century odyssey. *Annual Review of Clinical Psychology, 16,* 1-24.

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

Beauchaine, T. P., & McNulty, T. (2013). Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing psychopathology. Development and psychopathology, 25(4 0 2), 1505.

Blais, M. A., & Sinclair, S. J. (2018). *SPECTRA: Indices of Psychopathology*. Lutz, FL: PAR.

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality*. Amsterdam: Sociometric Research Foundation.

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biological Psychiatry.*

Boudreaux, M., South, S.C., & Oltmanns, T.F. (2019). Symptom-level analysis of DSM-IV/DSM-5 personality pathology in later life: Hierarchical structure and predictive validity across self- and informant ratings. *Journal of Abnormal Psychology, 128,* 365-384.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research (2nd edition)*. New York: Guilford.

Caspi, A., Houts, R. M., Ambler, A., Danese, A., Elliott, M. L., …, & Moffitt, T. E. (2020). Longitudinal assessment of mental disorders and comorbidities across four decades among participants in the Dunedin birth cohort Study. *JAMA Network Open, 3*(4), e203221.

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... & Moffitt, T. E. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders?. *Clinical Psychological Science, 2(2),* 119-137.

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry, 175,* 831-844.

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0. 2020.

Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods, 11,* 296-325.

Clark, L. A., Cuthbert, B. N., Lewis-Fernandéz, R., Narrow, W., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and RDoC. *Psychological Science in the Public Interest, 18(2),* 72-145.

Clark, L. A., Simms, L. J., Wu, K. D., & Casillas, A. (2014). *Schedule for Nonadaptive and Adaptive Personality—second edition*. Notre Dame: University of Notre Dame. Published by the first author.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100(3),* 316-336.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods, 19(2),* 300–315.

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., ... & Eaton, N. (2019). A hierarchical taxonomy of psychopathology can transform mental health research. *Perspectives on Psychological Science, 14(3),* 419-436.

Conway, C. C., & Simms, L. J. (2020). Maximizing the applied value of structural models of psychopathology: Introduction to a Special Issue of Personality and Mental Health. *Personality and Mental Health, 14,* 3-8.

Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S., & Zinbarg, R. E. (2009). What is an anxiety disorder?. *Depression and Anxiety, 26,* 1066-1085.

Cronbach, L. J., & Gleser, C. G. (1957). *Psychological tests and personnel decisions*. Urbana: University of Illinois Press.

Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), Decade of behavior. New methods for the analysis of change (p. 107–135). American Psychological Association.

Deary, I. J. (2012). Intelligence. *Annual Review of Psychology, 63,* 453-482.

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., ... & Baas, J. M. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety, 32(4),* 239-253.

Eaton, N. R. (2017). Advances in transdiagnostic psychopathology research: Introduction to the special issue. *Comprehensive Psychiatry, 79,* 1-3.

Eaton, N. R., Krueger, R. F., & Oltmanns, T. F. (2011). Aging and the structure and long-term stability of the internalizing spectrum of personality and psychopathology. Psychology and Aging, 26(4), 987.

Egan, R. P., & Smith, D. A. (in press). Marital quality and psychopathology: A meta-analytic investigation of general and specific affinity for mental disorders. *Psychological Bulletin*.

Forbes, M. K., Greene, A. L., Levin-Aspenson, H. F., Watts, A. L., Hallquist, M., Lahey, B. B., Markon, K. E., Patrick, C. J., Tackett, J. L., Waldman, I. D., Wright, A. G. C., Caspi, A., Ivanova, M., Kotov, R., Samuel, D. B., Eaton, N. R., & Krueger, R. F. (in press). A Comparison of the Reliability and Validity of the Predominant Models Used in Research on the Empirical Structure of Psychopathology. *Journal of Abnormal Psychology*. https://osf.io/fhp2r/

Forbes, M. K., Magson, N. R., & Rapee, R. M. (2020). Evidence that different types of peer victimization have equivalent associations with transdiagnostic psychopathology in adolescence. *Journal of Youth and Adolescence, 49,* 590-604.

Forbes, M., Rapee, R., & Krueger, R. (2019). Opportunities for the prevention of mental disorders by reducing general psychopathology in early childhood. *Behaviour Research and Therapy, 119,* 1-8.

Forbes, M. K., Tackett, J. L., Markon, K. E., & Krueger, R. F. (2016). Beyond comorbidity: Toward a dimensional and hierarchical approach to understanding psychopathology across the life span. Development and psychopathology, 28(4pt1), 971-986.

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR* D study. *Journal of Affective Disorders, 172,* 96-102.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. Journal of the American Academy of Child & Adolescent Psychiatry, 40(11), 1337-1345.

Gordon, J. A., & Redish, A. D. (2016). On the cusp. Current challenges and promises in psychiatry. In A. D. Redish & J. A. Gordon (Eds.), *Computational psychiatry: New perspectives on mental illness* (pp. 3-14). Cambridge, MA: MIT Press.

Hankin, B. L., Young, J. F., Abela, J. R., Smolen, A., Jenness, J. L., Gulley, L. D., ... & Oppenheimer, C. W. (2015). Depression from childhood into late adolescence: influence of gender, development, genetic susceptibility, and peer stress. Journal of abnormal psychology, 124(4), 803.

Hamlat, E.J., Young, J.F., & Hankin, B.L. (2019). Pubertal Timing as a Transdiagnostic Risk for Psychopathology in Youth. *Clinical Psychological Science, 7*, 411-429.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York, NY: Guilford Press.

Harter, S. (2012). Self-perception profile for adolescents: Manual and questionnaires. Denver, CO: Univeristy of Denver, Department of Psychology.

Hill, K. E., South, S. C., Egan, R. P., & Foti, D. (2019). Abnormal emotional reactivity in depression: Contrasting theoretical models using neurophysiological data. *Biological Psychology, 141,* 35-43.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. Experimental aging research, 18(3), 117-144.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). semTools: Useful tools for structural equation modeling. R package version 0.5-3. Retrieved from https://CRAN.R-project.org/package=semTools

Kline, R. B. (2015). *Principles and practice of structural equation modeling (4th edition).* New York: Guilford Press.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., … & Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126(4),* 454–477.

Kotov, R., Krueger, R. F., Watson, D., Cicero, D. C., Conway, C. C., DeYoung, C. G., Eaton, N. R., Forbes, M. K., Hallquist, M. N., Latzman, R. D., Mullins-Sweatt, S. N., Ruggero, C. J., Simms, L. J., Waldman, I. D., Waszczuk, M. A., & Wright, A. G. C. (under review). The Hierarchical Taxonomy of Psychopathology: A quantitative nosology based on consensus of evidence. Annual Review of Clinical Psychology

Kovacs, M. (1985). The Children's Depression Inventory (CDI). Psychopharmacology Bulletin, 21, 995–998.

Kozak, M. J., & Cuthbert, B. N. (2016). The NIMH research domain criteria initiative: background, issues, and pragmatics. *Psychophysiology, 53,* 286-297.

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42(9),* 1879-1890.

Krueger, R. F., Kotov, R., Watson, D., Forbes, M. K., Eaton, N. R., Ruggero, C. J., ... & Bagby, R. M. (2018). Progress in achieving quantitative classification of psychopathology. *World Psychiatry, 17(3),* 282-293.

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology, 116(4),* 645–666.

Krueger, R. F., Tackett, J. L., & MacDonald, A. (2016). Toward validation of a structural approach to conceptualizing psychopathology: A special section of the Journal of Abnormal Psychology. *Journal of Abnormal Psychology, 125(8),* 1023-1026.

Krueger, R. F., Watson, D., & Widiger, T. A. (2020). The vibrant intersection of personality and psychopathology research: A special issue of the Journal of Research in Personality. *Journal of Research in Personality, 84,* 1-3.

La Greca, A. M., & Harrison, H. M. (2005). Adolescent peer relations, friendships, and romantic relationships: Do they predict social anxiety and depression?. *Journal of clinical child and adolescent psychology, 34(1),* 49-61.

Laceulle, O. M., Vollebergh, W. A., & Ormel, J. (2015). The structure of psychopathology in adolescence: Replication of a general psychopathology factor in the TRAILS study. Clinical Psychological Science, 3(6), 850-860.

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology, 121,* 971-977.

Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). Validity and utility of the general factor of psychopathology. *World Psychiatry, 16,* 142-144.

Latzman, R.D., DeYoung, C.G., & The HiTOP Neurobiological Foundations Workgroup. (in press). Using empirically-derived dimensional phenotypes to accelerate clinical

neuroscience: The Hierarchical Taxonomy of Psychopathology (HiTOP) framework. *Neuropsychopharmacology*.

Lee, T. T. C., Sellbom, M., & Hopwood, C. J. (2017). Contemporary psychopathology assessment: Mapping major personality inventories onto empirical models of psychopathology. In S.C. Bowden (Ed.), *Neuropsychological Assessment in the Age of Evidence-Based Practice: Diagnostic and Treatment Evaluations* (pp. 64-94). New York: Oxford University Press.

Leyro, T. M., Zvolensky, M. J., & Bernstein, A. (2010). Distress tolerance and psychopathological symptoms and disorders: A review of the empirical literature among adults. *Psychological Bulletin, 136(4),* 576–600.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. Annual Review of Psychology, 60, 577–605.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130-149.

MacCallum, R.C., Wegener, D.T., Uchino, B.N., & Fabrigar, L.R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114(1),* 185–199.

Macatee, R. J., Capron, D. W., Guthrie, W., Schmidt, N. B., & Cougle, J. R. (2015). Distress tolerance and pathological worry: Tests of incremental and prospective relationships. *Behavior Therapy, 46(4),* 449-462.

Maples, J. L., Carter, N. T., Few, L. R., Crego, C., Gore, W. L., Samuel, D. B., … & Miller, J. D. (2015). Testing whether the DSM-5 personality disorder trait model can be measured

with a reduced set of items: An item response theory investigation of the Personality Inventory for DSM-5. *Psychological Assessment, 27(4),* 1195–1210.

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 554-565.

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88(1),* 139–157.

Michelini, G., Barch, D. M., Tian, Y., Watson, D., Klein, D. N., & Kotov, R. (2019). Delineating and validating higher-order dimensions of psychopathology in the Adolescent Brain Cognitive Development (ABCD) study. *Translational Psychiatry, 9,* 1-15.

Michelini, G., Palumbo, I. M., DeYoung, C. G., Latzman, R., & Kotov, R. (2020, June 10). Linking RDoC and HiTOP: A new interface for advancing psychiatric nosology and neuroscience. https://doi.org/10.31234/osf.io/ps7tc

Mineka, S., Watson, D. W., & Clark, L. A. (1998). Psychopathology: Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49,* 377-412.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural equation modeling, 9,* 599-620.

Nazarinia, R. R., Schumm, W. R., & White, J. M. (2009). Dimensionality and reliability of a modified version of Norton's 1983 Quality Marriage Index among expectant and new Canadian mothers. *Psychological Reports, 104(2),* 379-387.

Nelson, G. H., O'Hara, M. W., & Watson, D. (2018). National norms for the expanded version of the inventory of depression and anxiety symptoms (IDAS-II). *Journal of Clinical Psychology, 74(6),* 953-968.

Noordhof, A., Krueger, R. F., Ormel, J., Oldehinkel, A. J., & Hartman, C. A. (2015). Integrating autism-related symptoms into the dimensional internalizing and externalizing model of psychopathology. The TRAILS Study. *Journal of Abnormal Child Psychology, 43,* 577-587.

Olino, T. M., Bufferd, S. J., Dougherty, L. R., Dyson, M. W., Carlson, G. A., & Klein, D. N. (2018). The development of latent dimensions of psychopathology across early childhood: Stability of dimensions and moderators of change. Journal of abnormal child psychology, 46(7), 1373-1383.

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2020). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology.* Advance online publication. http://dx.doi.org/10.1037/pspp0000358

Pedhazur, E. J. (1997). *Multiple regression in behavioral research (3rd ed.)*. Orlando, FL: Harcourt Brace.

Perkins, E. R., Joyner, K. J., Patrick, C. J., Bartholow, B. D., Latzman, R. D., DeYoung, C. G., ... & Docherty, A. R. (2020). Neurobiology and the Hierarchical Taxonomy of Psychopathology: progress toward ontogenetically informed and clinically useful nosology. *Dialogues in Clinical Neuroscience, 22,* 51-63.

Perkins, E. R., Latzman, R. D., & Patrick, C. J. (2020). Interfacing neural constructs with the Hierarchical Taxonomy of Psychopathology: 'Why' and 'how'. *Personality and Mental Health, 14,* 106-122.

Prinstein, M. J., & Giletta, M. (2016). Peer relations and developmental psychopathology. In D. Cicchetti (Ed.), *Developmental Psychopathology* (pp. 1–53). Hoboken, NJ: Wiley. https://doi.org/10.1002/9781119125556.devpsy112

Rapee, R. M., Oar, E. L., Johnco, C. J., Forbes, M. K., Fardouly, J., Magson, N. R., & Richardson, C. E. (2019). Adolescent development and risk for the onset of social-emotional disorders: A review and conceptual model. *Behaviour research and therapy*, *123*. https://doi.org/10.1016/j.brat.2019.103501

Rodriguez-Seijas, C., Stohl, M., Hasin, D. S., & Eaton, N. R. (2015). Transdiagnostic factors and mediation of the relationship between perceived racial discrimination and mental disorders. *Journal of the American Medical Association Psychiatry, 72,* 706-713.

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E., … & Zimmermann, J. (2019). Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. *Journal of Consulting and Clinical Psychology, 87(12),* 1069–1084.

Sellbom, M., Carragher, N., Sunderland, M., Calear, A. L., & Batterham, P. J. (2020). The role of maladaptive personality domains across multiple levels of the HiTOP structure. *Personality and Mental Health, 14,* 30-50.

Simms, L., Wright, A. G. C., Cicero, D., Kotov, R., Mullins-Sweatt, S. N., …, & Zimmermann, J. (under review). Development of measures for the Hierarchical Taxonomy of Psychopathology (HiTOP): A collaborative scale development project. *Assessment*.

Simons, J. S., & Gaher, R. M. (2005). The Distress Tolerance Scale: Development and validation of a self-report measure. *Motivation and Emotion, 29(2),* 83-102.

Smith, G. T., McCarthy, D. M., & Zapolski, T. C. B. (2009). On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. *Psychological Assessment, 21,* 272-284.

Snyder, H., Friedman, N., & Hankin, B.L. (2019). Transdiagnostic mechanisms of psychopathology in youth: Executive functions, dependent stress, and rumination. *Cognitive Therapy and Research, 43,* 834-851.

Snyder, H. R., Young, J. F., & Hankin, B. L. (2017). Strong homotypic continuity in common psychopathology-, internalizing-, and externalizing-specific factors over time in adolescents. Clinical Psychological Science, 5(1), 98-110.

South, S. C. (in press). Pathology in relationships. *Annual Review of Clinical Psychology*.

South, S. C., & Krueger, R. F. (2008). Marital quality moderates genetic and environmental influences on the internalizing spectrum. *Journal of Abnormal Psychology, 117(4),* 826–837.

South, S. C., Krueger, R. F., & Iacono, W. G. (2011). Understanding general and specific connections between psychopathology and marital distress: A model based approach. *Journal of Abnormal Psychology, 120,* 935-947.

Stasik-O'Brien, S. M., Brock, R. L., Chmielewski, M., Naragon-Gainey, K., Koffel, E., McDade-Montez, E., ... & Watson, D. (2019). Clinical utility of the inventory of depression and anxiety symptoms (IDAS). *Assessment, 26(5),* 944-960.

Steinberg, L., & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology, 52(1),* 83–110. https://doi.org/10.1146/annurev.psych.52.1.83.

Sunderland, M., Slade, T., Carragher, N., Batterham, P., & Buchan, H. (2013). Age-related differences in internalizing psychopathology amongst the Australian general population. Journal of Abnormal Psychology, 122(4), 1010–1020.

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods, 24*(5), 637-657. http://dx.doi.org/10.1037/met0000210

Vollebergh, W. A., Iedema, J., Bijl, R. V., de Graaf, R., Smit, F., & Ormel, J. (2001). The structure and stability of common mental disorders: the NEMESIS study. Archives of general psychiatry, 58(6), 597-603.

Waszczuk, M. A., Eaton, N. R., Krueger, R. F., Shackman, A. J., Waldman, I. D., Zald, D. H., Lahey, B. B., Patrick, C. J., Conway, C. C., Ormel, J., Hyman, S. E., Fried, E. I., Forbes, M. K., Docherty, A. R., Althoff, R. R., Bach, B., Chmielewski, M., DeYoung, C. G., Forbush, K. T., . . . Kotov, R. (2020). Redefining phenotypes to advance psychiatric genetics: Implications from hierarchical taxonomy of psychopathology. *Journal of Abnormal Psychology, 129,* 143-161.

Waszczuk, M. A., Kotov, R., Ruggero, C., Gamez, W., & Watson, D. (2017). Hierarchical structure of emotional disorders: From individual symptoms to the spectrum. *Journal of Abnormal Psychology, 126,* 613-634.

Waszczuk, M. A., Zimmerman, M., Ruggero, C., Li, K., MacNamara, A., Weinberg, A., ... & Kotov, R. (2017). What do clinicians treat: Diagnoses or symptoms? The incremental validity of a symptom-based, dimensional characterization of emotional disorders in predicting medication prescription patterns. *Comprehensive Psychiatry, 79*, 80-88.

Watson, D. (2009). Differentiating the mood and anxiety disorders: A quadripartite model.

Annual Review of Clinical Psychology, 5, 221-247.

Watson, D., Clark, L. A., Chmielewski, M., & Kotov, R. (2013). The value of suppressor effects in explicating the construct validity of symptom measures. *Psychological Assessment, 25,* 929-941.

Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., ... & Ruggero, C. J. (2012). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment, 19,* 399-420.

Watson, D., O'Hara, M. W., Simms, L. J., Kotov, R., Chmielewski, M., McDade-Montez, E. A., Gamez, W., & Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological Assessment, 19(3),* 253-268.

Watts, A. L., Poore, H. E., & Waldman, I. D. (2019). Riskier tests of the validity of the bifactor model of psychopathology. *Clinical Psychological Science, 7,* 1285-1303.

Widiger, T. A., Lynam, D. R., Miller, J. D., & Oltmanns, T. F. (2012). Measures to assess maladaptive variants of the five-factor model. *Journal of Personality Assessment, 94(5),* 450-455.

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and statistical manual of mental disorders--fifth edition. *Journal of Abnormal Psychology, 114(4),* 494–504.

World Health Organization (WHO). (2018). *International classification of diseases for mortality and morbidity statistics* (11th revision). Geneva, Switzerland: Author.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73,* 913-934.

Wong, K. T. K., & Forbes, M. K. (2020). Evidence that relationship quality, social support, and SSRI use do not account for the shared underlying relationships among symptoms of depression, anxiety, and female sexual dysfunction. *The Journal of Sex Research, 57,* 365-374.

Table 1. Regression results for distress tolerance outcomes for Tutorial 1.

	Distress Tolerance Scale							
	Simple Regression				Multiple Regression			
Predictor	*b*	*SE*	*p*	*β*	*b*	*SE*	*p*	*β*
Dysphoria	-7.38	0.83	.000	-.60	—	—	—	—
Lassitude	-5.11	0.94	.000	-.41	-1.06	1.12	.344	-.08
Insomnia	-3.10	1.00	.002	-.25	-0.85	0.96	.378	-.07
Appetite loss	-4.53	0.96	.000	-.37	-0.62	1.14	.586	-.05
Appetite gain	-2.46	1.01	.016	-.20	-1.46	0.93	.117	-.12
Well-being	4.82	0.95	.000	.39	2.21	1.05	.037	.18
Ill-temper	-3.31	1.00	.001	-.27	-1.41	0.94	.135	-.11
Traumatic intrusions	-5.01	0.94	.000	-.41	-0.48	1.26	.700	-.04
Social anxiety	-4.79	0.95	.000	-.39	-1.04	1.10	.345	-.08
Panic	-5.41	0.93	.000	-.44	-2.15	1.35	.114	-.17
Suicidality	-4.21	0.97	.000	-.34	-1.18	1.03	.253	-.09

Note. All predictors were standardized prior to analysis.

Table 2. Comparing the total and direct effects of relationship quality on symptom dimensions at different levels of the psychopathology hierarchy for Tutorial 2.

Psychopathology Outcome	Total Effect (*SE*)	Direct Effect (*SE*)	Proportion accounted for by higher-order factor(s)[a]
Depression	**-.37 (.04)**	**-.13 (.03)**	65%
Generalized Anxiety	**-.26 (.04)**	**.14 (.04)**	100%
Social Anxiety	**-.24 (.04)**	.01 (.04)	100%
Panic	**-.28 (.07)**	-.07 (.07)	77%
Obsessions and compulsions	**-.24 (.04)**	.02 (.04)	100%
Distress	**-.36 (.04)**	-.05 (.04)	86%
Fear	**-.32 (.05)**	.05 (.04)	100%
Internalizing	**-.37 (.04)**	**-.37 (.04)**[b]	-

Note. All effects are fully standardized estimates from Mplus version 8.0. Bolded effects are statistically significant ($p < .001$). Total effects represent the expected change in psychopathology outcomes (in standard deviation units) per standard deviation increase in relationship quality. Direct effects represent the fully standardized effect of relationship quality on psychopathology outcomes, above and beyond any higher-order dimensions (path d in Figure 3). The indirect effect can be computed by subtracting the direct effect from the total effect. See the main text for more detail.

[a] Some direct effects were opposite in sign to the corresponding total effects. Consider generalized anxiety, for example. Its "zero-order" association with relationship quality was -.26 (total effect), whereas the partial association (after "partialling out" variance generalized anxiety shared with other symptom dimensions; direct effect) was .14. This pattern implies that relationship quality is *inversely* related to generalized anxiety on a bivariate level, but *positively* related to the part of generalized anxiety that does not overlap with other symptom dimensions (cf. suppression; Watson, Clark, Chmielewski, & Kotov, 2013). When direct effects had opposite signs to the relevant total effects, we indicated that the proportion of the total effect accounted for by higher-order factors (final column in the table) was 100%. In other words, these positive direct effects did not augment the overall (total) negative effect of relationship quality on the psychopathology outcome.

[b] The total and direct effects for internalizing are equivalent because internalizing is the highest level of the psychopathology hierarchy here. There are no broader (i.e., superordinate) dimensions through which relationship quality could have (indirectly) influenced internalizing. Because there are no superordinate (i.e., higher-order) factors, relative to internalizing, in the model, the final column in this table is left blank.

Figure Captions

Figure 1. Working Hierarchical Taxonomy of Psychopathology (HiTOP) consortium model. Constructs higher in the figure are broader and more general, whereas constructs lower in the figure are narrower and more specific. Dashed lines denote provisional elements requiring further study. At the lowest level of the hierarchy (i.e., traits and symptom components), for heuristic purposes, conceptually related signs and symptoms (e.g., phobia) are indicated in bold, with specific manifestations indicated in parentheses. ADHD = attention-deficit/hyperactivity disorder; BPD = bipolar disorder; GAD = generalized anxiety disorder; HiTOP = Hierarchical Taxonomy of Psychopathology; IED = intermittent explosive disorder; MDD = major depressive disorder; OCD = obsessive-compulsive disorder; ODD = oppositional defiant disorder; SAD = separation anxiety disorder; PD = personality disorder; PTSD = posttraumatic stress disorder.

Figure 2. Direct and indirect effects of relationship quality on emotional-disorder symptoms. This diagram illustrates the regression of the internalizing factor and depression symptoms on relationship quality. The indirect effect of relationship quality on depression via internalizing is represented by the path a > b > c. In contrast, the direct effect of relationship quality on depression is represented by path d. These path names mirror the labeling scheme we use in our data analysis code in the online supplement. For clarity of presentation, here we omit the indirect pathway from relationship quality to depression via the distress factor. The total effect of relationship quality on depression is the sum of the direct effect (relationship quality > depression), the indirect effect via internalizing (relationship quality > internalizing > distress > depression), and the indirect effect via distress (relationship quality > distress > depression). Regarding notation, dep = depression; gad = generalized anxiety; pan = panic; soc = social phobia; obs = obsessions and compulsions. Rectangles represent observed variables and ellipses represent latent factors. Short, single-headed arrows pointing to distress, fear, and the factor indicators reflect error variances. Single-headed arrows from internalizing to distress and fear represent second-order factor loadings; single-headed arrows from distress and fear to factor indicators are first-order factor loadings.

Figure 3. Cross-lagged panel model of internalizing and social competence over three years. All effect sizes are fully standardized (see Table S6 for full results). All factor loadings were statistically significant at the .001 alpha level. Rectangles and circles represent observed and latent variables, respectively. Dashed lines denote statistically nonsignificant paths. *com* = social competence; *int* = internalizing; CDI = Children's Depression Inventory; MASC = Manifest Anxiety Scale for Children; SDQ = Strengths and Difficulties Questionnaire. *** $p < .001$.

Figure 1

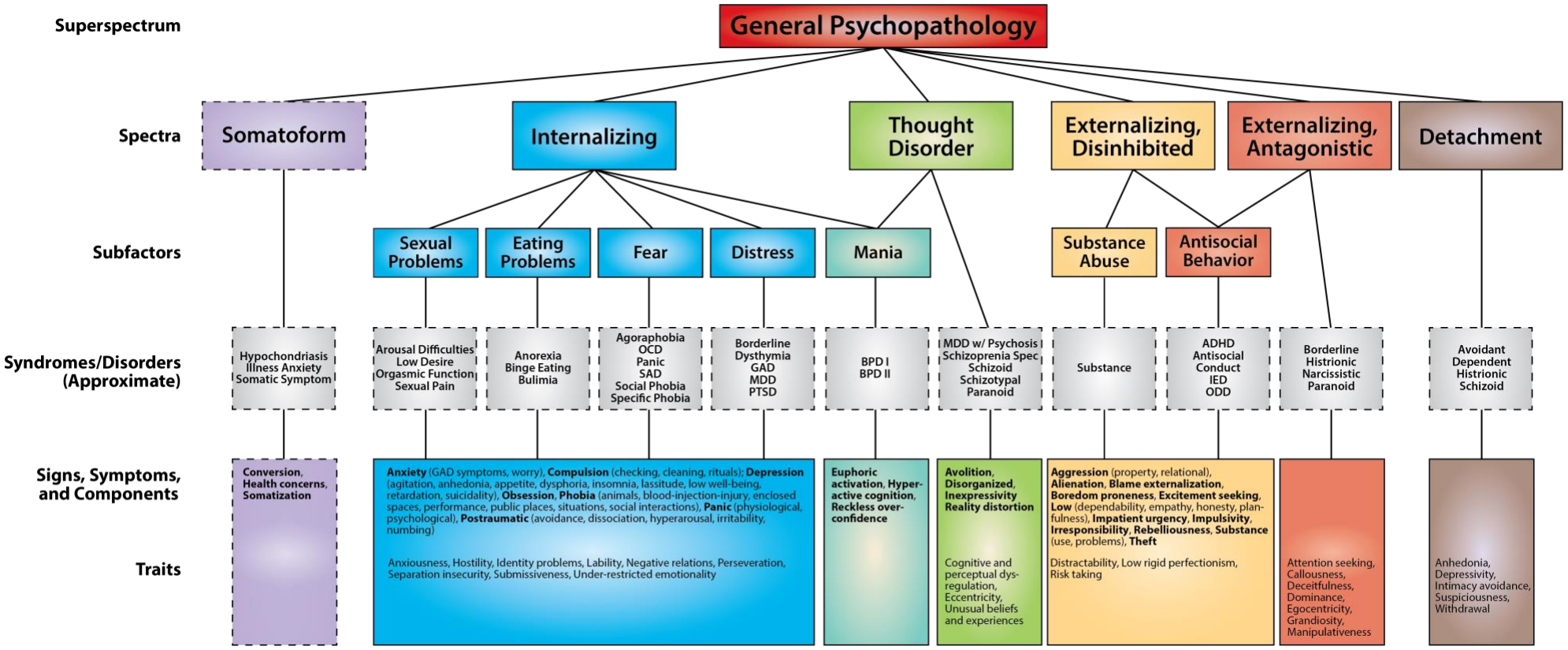

Superspectrum
General Psychopathology
Spectra
Somatoform
Internalizing
Thought Disorder
Externalizing, Disinhibited
Externalizing, Antagonistic
Detachment
Subfactors
Sexual Problems
Eating Problems
Fear
Distress
Mania
Substance Abuse
Antisocial Behavior
Syndromes/Disorders (Approximate)
Hypochondriasis
Illness Anxiety
Somatic Symptom
Arousal Difficulties
Low Desire
Orgasmic Function
Sexual Pain
Anorexia
Binge Eating
Bulimia
Agoraphobia
OCD
Panic
SAD
Social Phobia
Specific Phobia
Borderline
Dysthymia
GAD
MDD
PTSD
BPD I
BPD II
MDD w/ Psychosis
Schizoprenia Spec
Schizoid
Schizotypal
Paranoid
Substance
ADHD
Antisocial
Conduct
IED
ODD
Borderline
Histrionic
Narcissistic
Paranoid
Avoidant
Dependent
Histrionic
Schizoid
Signs, Symptoms, and Components
Traits
Conversion, Health concerns, Somatization
Anxiety (GAD symptoms, worry), Compulsion (checking, cleaning, rituals); Depression (agitation, anhedonia, appetite, dysphoria, insomnia, lassitude, low well-being, retardation, suicidality), Obsession, Phobia (animals, blood-injection-injury, enclosed spaces, performance, public places, situations, social interactions), Panic (physiological, psychological), Postraumatic (avoidance, dissociation, hyperarousal, irritability, numbing)
Anxiousness, Hostility, Identity problems, Lability, Negative relations, Perseveration, Separation insecurity, Submissiveness, Under-restricted emotionality
Euphoric activation, Hyper-active cognition, Reckless over-confidence
Avolition, Disorganized, Inexpressivity
Reality distortion
Cognitive and perceptual dys-regulation, Eccentricity, Unusual beliefs and experiences
Aggression (property, relational), Alienation, Blame externalization, Boredom proneness, Excitement seeking, Low (dependability, empathy, honesty, plan-fulness), Impatient urgency, Impulsivity, Irresponsibility, Rebelliousness, Substance (use, problems), Theft
Distractability, Low rigid perfectionism, Risk taking
Attention seeking, Callousness, Deceitfulness, Dominance, Egocentricity, Grandiosity, Manipulativeness
Anhedonia, Depressivity, Intimacy avoidance, Suspiciousness, Withdrawal

Figure 2

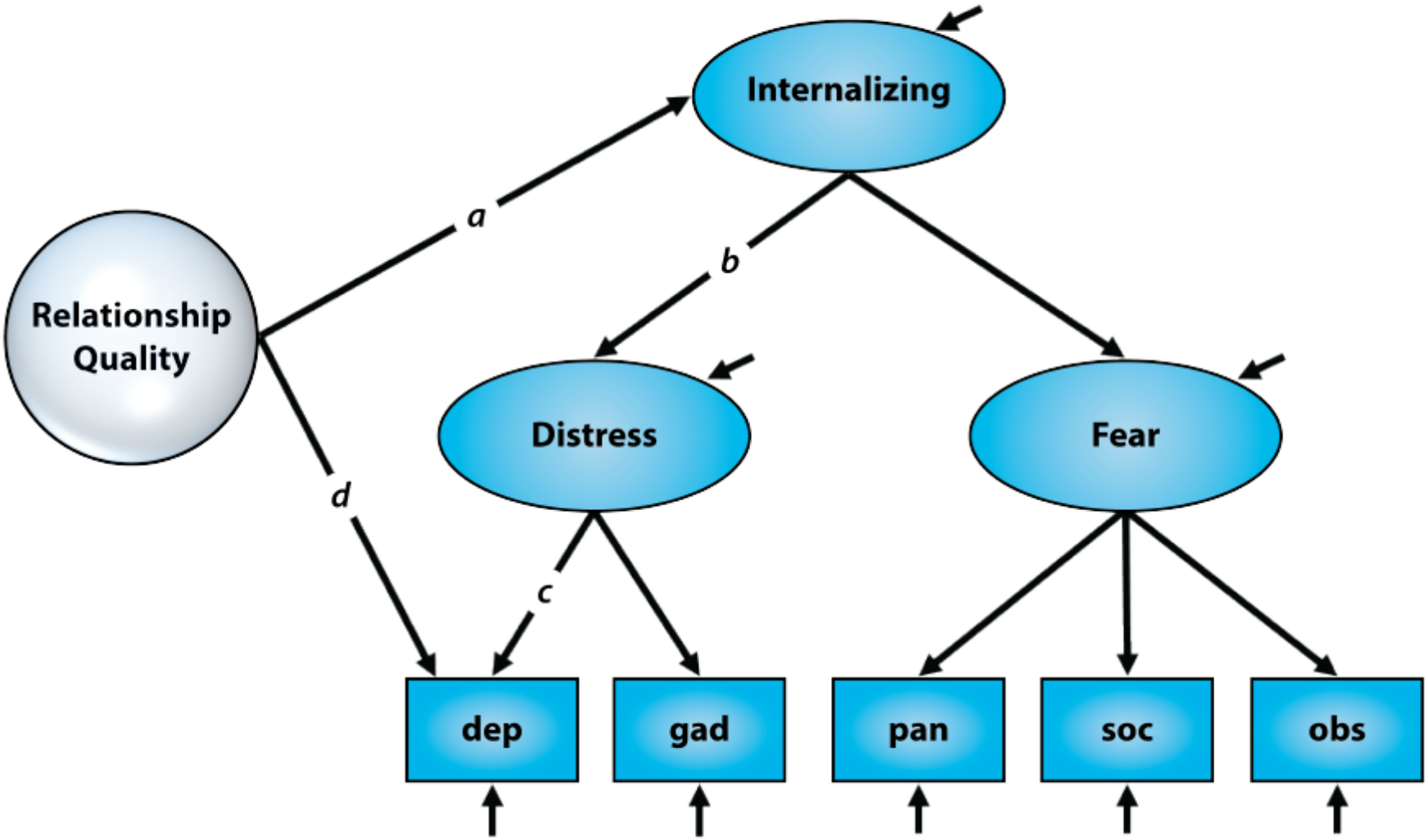
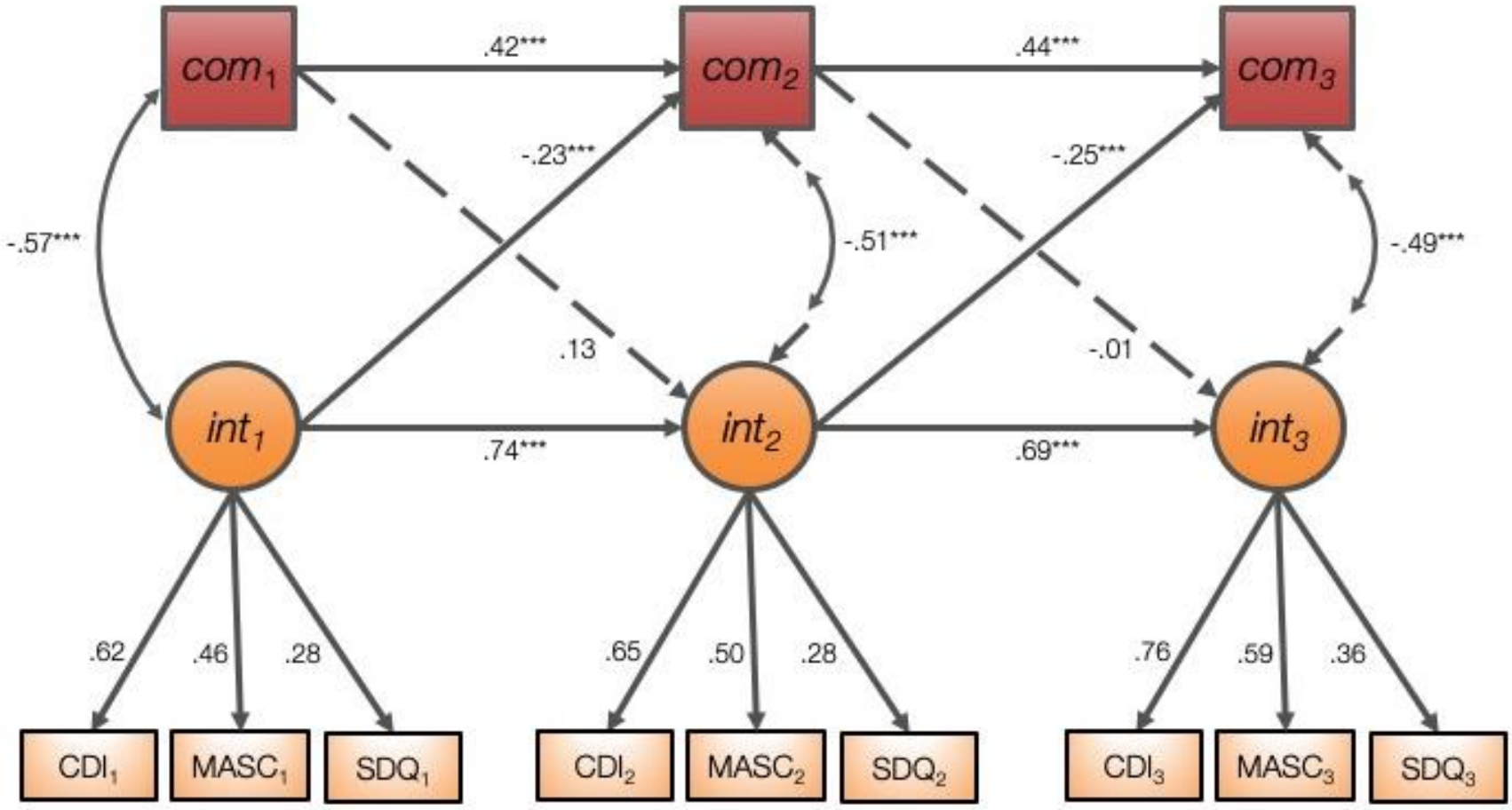

Internalizing
Relationship Quality
a
b
d
Distress
Fear
c
dep
gad
pan
soc
obs

Figure 3


com1
com2
com3
int1
int2
int3
.42***
.44***
-.23***
-.25***
-.57***
-.51***
-.49***
.13
-.01
.74***
.69***
.62
.46
.28
.65
.50
.28
.76
.59
.36
CDI1
MASC1
SDQ1
CDI2
MASC2
SDQ2
CDI3
MASC3
SDQ3