# Disquiet in nosology: A primer on an emerging, empirically-based approach to classifying mental illness and implications for training

Camilo J. Ruggero, Ph.D. [a], Jennifer L. Callahan, Ph.D. [a], Allison Dornbach-Bender, M.S. [a], Jennifer L. Tackett, Ph.D. [b], Roman Kotov, Ph.D. [c]

[a] Department of Psychology, University of North Texas, Denton TX

[b] Department of Psychology, Northwestern University, Evanston, IL

[c] Department of Psychiatry, Stony Brook University, Stony Brook, NY



Corresponding author: Camilo J. Ruggero, Ph.D., University of North Texas, 1155 Union Circle #311280, Denton, TX 76203, Phone: (940) 565-3291, Fax: (940) 565-4682, Email: Camilo.Ruggero@unt.edu

## I. Prevailing mental health nosologies: a caution

Paul Meehl (1986) warned more than 30 years ago of a "scientific malignancy" worth recalling: the tendency by some to reify diagnoses, as though the criteria that operationalize a disorder in the *Diagnostic and Statistical Manual of Mental Disorders (DSM;* APA, 2013) describe its essence. Diagnoses, instead, are open constructs.[1] Most of us, when pressed, easily acknowledge the difference. The core motivation behind the National Institute of Mental Health's Research Domain Criteria (Cuthbert & Insel, 2013) underscores this point. Yet when not pressed, too often the criteria can slip into becoming the disorder. It would be unfair to blame *DSM* for this habit (cf Kraemer, Kupfer, Clarke, Narrow, & Regier, 2012), yet its operationalization of criteria risks making us forget that articulating a useful mental health nosology remains ongoing.

Prevailing classification approaches have other problems. Disorders are presumed distinct, yet the predominance of comorbidity raises obvious questions about the validity of their borders (e.g., Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Kessler, Chiu, Demler, & Walters, 2005; Ormel et al., 2015; Teesson, Slade, & Mills, 2009). Or, categories can have marked heterogeneity, such that two individuals with the same diagnosis have entirely different sets of symptoms (Clark, Watson, & Reynolds, 1995; Hasler, Drevets, Manji, & Charney, 2004; Zimmerman, Ellison, Young, Chelminski, & Dalrymple, 2015). Reliability is often too low (Chmielewski, Clark, Bagby, & Watson, 2015; Regier et al., 2013), and evidence overwhelmingly suggests psychopathology falls along a continuum, with no clear zones of rarity

[1] This can be debated, of course (see Wakefield, 2004).

(Wright et al., 2013). Finally, it is not always clear from surveys how clincially useful clinicians find the prevailing nosology beyond its relevance for billing (First et al., 2018).

Despite these concerns, nosology remains foundational for anyone whose work intersects with mental health (Blashfield & Burgess, 2007). At minimum, it gives us a *lingua franca* to talk about symptoms and how they present. But ideally it would do so much more: it would guide our treatments, forecast the course of illness, and create a foundation for research into the causes of illness (Mullins-Sweatt, Lengel & DeShong, 2016). For students in training, *DSM's* lexicon, and the assumptions behind it, get woven into their curriculum and shape conceptualizations of psychopathology (e.g., Amazon ranks *DSM* second in psychology reference books, only behind the American Psychological Association's style manual).

**II. Next generation approach**

*DSM's* hegemony over classification has overshadowed an accelerating body of research happening in the wings of mental health, largely driven by psychologists: quantitative nosology. At its core, this approach creates a data-driven, empirically-based classification. It starts with diverse arrays of highly homogenous signs and symptoms of mental health problems (e.g., dysphoric mood). Statistical procedures like factor analyses and hierarchical agglomerative clustering are then used to organize elements into increasingly more heterogonous, higher order constructs based on patterns of association.

This method is hardly new: Thomas Moore in the 1930's analyzed the intercorrelations among 32 signs and symptoms related to psychosis to understand how they could be more parsimoniously grouped into higher order factors. Many others, notably Achenbach and colleagues (Achenbach, 1966; Achenbach, Ivanova, & Rescorla, 2017), followed suit with increasing sophistication and precision (Kotov, 2016).

The most recent large-scale effort in this movement toward empirically based classification emerged in the spring of 2015. Forty scholars working in the area of quantitative nosology started a consortium (now close to 100 members) devoted to articulating an empirically-based quantitative nosology of mental illness. Their initial proposed model - the Hierarchical Taxonomy of Psychopathology (HiTOP; Kotov et al., 2017) - provides a marked departure from nosology systems like *DSM*.

**III. HiTOP: a primer**

HiTOP's empirically-based model remains a work in progress (remember Meehl's admonitions!) and the consortium is actively working to revise the model as new evidence emerges (Krueger et al., 2018) but major, replicated contours of this nosology are already clear. The model is hierarchical, with homogenous signs, symptoms and traits at the bottom. There are over 100 of these dimensions, and they consist of symptom components, such as insomnia, and traits, such as submissiveness. These are organized into higher order components that are increasingly broad until one reaches what is called the spectra level – of which there are six (i.e., Internalizing, Somatoform, Thought Disorder, Disinhibited Externalizing, Antagonistic Externalizing and Detachment). Above this, one can aggregate higher all the way up to a general factor (i.e., so-called "p-factor;" Caspi et al., 2014). Figure 1 presents portions of the model, reprinted and revised with permission.

How does this differ from the *DSM*? With traditional nosology, symptoms related to depression, generalized anxiety disorder, and social phobia, to take one example, constitute three putatively distinct categories of mental illness. In contrast, with HiTOP they all fall under the rubric of an internalizing spectrum. A provider can focus on this higher level spectrum, recognizing that all three syndromes share elements. Or, one can cascade down the model, with

for example depression and generalized anxiety symptoms coalescing under a "distress" subfactor whereas social phobia hewing more closely to a "fear" subfactor. Or one can cascade even further down, focusing on highly homogenous symptoms or traits, such as suicidality.

Unlike *DSM*, HiTOP does not delineate a "onse size fits all" boundary between "illness" and "not illness," a feature supported by years of taxometric research (Haslam, Holland & Kuppens, 2012). Rather, clinical decisions are guided by ranges of severity on each dimension of the model. Until work validates these in different populations, they can remain statistical (e.g., 2 *SD* below the mean), such as with intelligence testing, or can be tailored to the needs and resources available within a given setting or population. Kotov et al. (2017) reviews evidence supporting the model, while Ruggero et al. (2018) provides a description of its integration into clinical care.



**IV. HiTOP may advance research and treatment**

HiTOP proposes to accelerate mental health research (Conway et al., 2019). Use of continuous dimensions, as opposed to categories, has well-known benefits for statistical power of research to detect effects (Cohen, 1983). Compared to categorical phenotypes, dimensional ones double the power to predict a variety of clinical outcomes (Kotov et al, 2019) and produce more "hits" in genetic research (Otowa et al., 2016), for example.

But the hierarchical structure in and of itself provides a novel framework for pursing pathophysiologies. Mechanisms, or outcomes, may operate at different levels of this mental illness hierarchy, from broad and diffuse effects to more narrow and specific ones. HiTOP's hierarchy provides one map to different levels that may be relevant, and at minimum new phenotypic targets on which to test proposed mechanisms. Already, work in genetics,

neurobiology, and psychosocial contexts point to how recent findings in these fields may better align with models like HiTOP compared to traditional nosology (Conway et al., 2019).

HiTOP also proposes potentially greater clinically utility (Ruggero et al., 2018). Dimensions are more reliable than traditional categories (e.g., 15% increased reliability in meta-analyses; Markon, Chmielewski, & Miller, 2011) and may be preferred over categories by clinicians (Morey, Skodol, & Oldham, 2014). Moreoever, HiTOP higher-level spectra may have increased prognostic power, for example predicting suicide attempts, future psychopathology and other clinical outcomes more than disorder-specific variation alone (Eaton et al., 2013; Kim & Eaton, 2015). HiTOP may also better align with treatment planning. Early evidence suggests clinician prescribing practices track more closely to a HiTOP-based model compared to a *DSM* one (Waszczuk et al., 2017). Similarly, emerging trandiagnsotic approaches to the treatment of mental health (e.g., Barlow et al., 2017) align well with HiTOP's conceptualization of upper level spectra that share features, and potential etiologies. Finally, HiTOP provides flexibility to adapt clinical ranges based on their purpose, rather than requiring one-size-fits-all cutoffs common to *DSM,* removing from nosology their reification that are not empirically based (e.g. five of nine symptoms because five is more than half). None of these advantages guarantee HiTOP's clinical utility, but they provide impetus for testing its utility and tackling the major challenge of training students on this new nosology.



**V. Training implications**

A caution against casually introducing any new nosology, particularly one based on dimensions, would be its implications and cost for training given the major investment already made in the use of *DSM* (First et al., 2005). Although a major concern for fields less accustomed to dimensional models (e.g., psychiatry), students in psychology are already well-trained in

working with conceptualizations and measures consistent with HiTOP (e.g., MMPI-2-RF, NEO-PI-3, PAI; Ben-Porath & Tellegen, 2008/2011; Costa & McCrae, 2010; Morey, 2007). Nevertheless, a shift toward HiTOP would impact courses related to foundational knowledge (i.e., psychopathology, assessment, and treatment planning), as well as functional competencies in the application of HiTOP via practicum and internship or residency experiences.

We field tested training in HiTOP at one of the author's (JLC) own universities to better appreciate the feasibility of weaving HiTOP into foundational parts of a curriculum. An assessment instructor (JLC) spoke with two members from the HiTOP consortium (RK and CJR) about the model. They provided training material, including slides for instruction. The instructor then developed curricular components for the three foundational knowledge areas. During the psychopathology component, the HiTOP model was overviewed in class after introduction of *DSM*. During the structured interviewing component, challenges of a *DSM* approach to assessment and case conceptualization were presented, including concerns about reliability, heterogeneity, and comorbidity. The HiTOP model was presented as an emerging alternative that resolved some of these problems, although remained untested with respect to its clinical utility. The lecture component concluded by overviewing a list of measures routinely taught in the course and used in practicum that are consistent with a HiTOP approach to case conceptualization. Finally, during treatment planning instruction, the HiTOP model was briefly reviewed, again drawing some content from the expert slides, before engaging in hypothetical clinical decision making exercises (e.g., using the HiTOP framework to identify the salient spectra that will become the focus of a transdiagnostic treatment; e.g., Barlow et al., 2017; Lundhal, Kunz, Brownell, Tollefson, & Burke, 2010).

Integration of HiTOP into these training components was seamless from the instructor's perspective and end-of-course satisfaction evaluations suggest the material was well-received by students. Foundational HiTOP knowledge was assessed as part of the midterm exam in the assessment course with all students meeting the threshold for at least adequate accuracy (70% or greater). Sequencing of HiTOP's introduction (first psychopathology, then assessment, and finally treatment-planning) flowed intuitively and was consistent with the larger curriculum. Given that the model includes many *DSM*-like constructs, albeit broken into smaller (symptom component) or larger (spectra) units in a hierarchical fashion, it was feasible to teach students the *DSM* categories for practical and perhaps temporary purposes, while familiarizing them as well with evidence-based hierarchical models.

Finally, it is common for students to learn how to apply cut scores along recognized continua, such as with IQ or use of T-scores common to many measures. Thus, students were taught to think about diagnostic cut scores for psychopathology diagnosis in the same way: diagnostic thresholds are indicators not of people who can be classified as qualitatively different from the healthy, but of relative severity on continua that suggest varying need for treatment. These experiences remain anecdotal, but they demonstrate the feasibility of weaving HiTOP training into existing psychology program curriculums. Importantly, this exercise found that HiTOP training could be integrated without major cost (from additional texts or new measures) and without radical changes to the core curriculum.



## VI. Conclusions

How we classify mental illness is foundational for psychologists, carrying profound implications for the research and treatment of mental illness, as well as training of future psychologists. Prevailing approaches lack the empirical support often called for (Krueger et al.,

2018) and suffer shortcomings, including reification, less than desired reliability, and questions about the validity of proposed categories. Quantitative nosology generally, and HiTOP as the latest synthesis of these models in particular, offers a departure from prevailing nosologies, with arguably more empirical support. Dimensions, not categories, are organized hierarchically. This new model's flexibility provides novel targets and a powerful framework for research, and may better align with treatment. Training remains a challenge for the broader mental health field, but HiTOP can already be integrated intuitively into psychology training curriculum.

**Disclosure Statement**: The authors have no funding or conflicts of interest to disclose.

## VII. References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.).* Arlington, VA: American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: a factor analytic study. *Psychological Monographs, 80*(7), 1-37. https://doi.org/10.1037/h0093906

Achenbach, T. M., Ivanova, M. Y., & Rescorla, L. A. (2017). Empirically based assessment and taxonomy of psychopathology for ages 1½-90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry, 79*, 4-18. https://doi.org/10.1016/j.comppsych.2017.03.006

Barlow, D. H., Farchione, T. J., Bullis, J. R., Gallagher, M. W., Murray-Latin, H., Sauer-Zavala, S., … & Cassiello-Robbins, C. (2017). The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders Compared With Diagnosis-Specific Protocols for Anxiety Disorders: A Randomized Clinical Trial. *JAMA Psychiatry. 74*(9), 875-884. https://doi.org/10.1001/jamapsychiatry.2017.2164

Blashfield, R. K., & Burgess, D. R. (2007). Classification provides an essential basis for organizing mental disorders. In S. O. Lilienfeld & W. T. O'Donohue (Eds.). *The Great Ideas Of Clinical Science: 17 Principles That Every Mental Health Professional Should Understand* (93–117). New York, NY: Routledge/Taylor & Francis.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical

sample. *Journal of Abnormal Psychology*, *110*(4), 585. https://doi.org/10.1037/0021-843X.110.4.585

Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2-Restructured Form): Manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, C. 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. https://doi.org/10.1177/2167702613497473

Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *Journal of Abnormal Psychology, 124*(3), 764-769. https://doi.org/10.1037/abn0000069

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46*(1), 121-153. https://doi.org/10.1146/annurev.ps.46.020195.001005

Cohen J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*, 249-253.

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., . . . & Eaton, N. R. (2019). A Hierarchical Taxonomy of Psychopathology can reform mental health research. *Perspectives on Psychological Science*. Advance online publication. https://doi.org/10.1177/1745691618810696

Costa, P. T., & McCrae, R. R. (2010). The NEO Personality Inventory: 3. Odessa, FL: *Psychological Assessment Resources*.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11*, 126–134. http://dx.doi.org/10.1186/1741-7015-11-126

Eaton, N. R., Krueger, R. F., Markon, K. E., Keyes, K. M., Skodol, A. E., Wall, M., … & Grant, B. F. (2013). The structure and predictive validity of the internalizing disorders. *Journal of Abnormal Psychology, 122*(1), 86-92. https://doi.org/10.1037/a0029598

First, M. B. (2005). Clinical utility: a prerequisite for the adoption of a dimensional approach in DSM. *Journal of Abnormal Psychology, 114*(4), 560-564. https://doi.org/10.1037/0021-843X.114.4.560

First, M. B., Rebello, T. J., Keeley, J. W., Bhargava, R., Dai, Y., Kulygina, M., ... & Reed, G. M. (2018). Do mental health professionals use diagnostic classifications the way we think they do? A global survey. *World Psychiatry, 17*(2), 187-195. https://doi.org/10.1002/wps.20525

Haslam, N., Holland, E., & Kuppens, P. (2012). Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychological Medicine, 42*(5), 903-920. https://doi.org/10.1017/S0033291711001966

Hasler, G., Drevets, W. C., Manji, H. K., & Charney, D. S. (2004). Discovering endophenotypes for major depression. *Neuropsychopharmacology, 29*(10), 1765-1781. https://doi.org/10.1038/sj.npp.1300506

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 617-627. https://doi.org/10.1001/archpsyc.62.6.617

Kim, H., & Eaton, N. R. (2015). The hierarchical structure of common mental disorders: Connecting multiple levels of comorbidity, bifactor models, and predictive validity. *Journal of Abnormal Psychology, 124*(4), 1064-1078. https://doi.org/10.1037/abn0000113

Kotov, R., Barch, D.M., Barlow, D. H., Carpenter, W., Hasin, D.S., Heckers, S.H., Krueger, R.F., Miller, G.A., Ruggero, C.J., Watson, D., Zald, D.H. (2019). Why quantitative nosology is the best investment in mental health today. Manuscript submitted for publication.

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. https://doi.org/10.1037/abn0000258

Kotov, R. (2016). The quantitative classification of mental illness: Emerging solution to boundary problems. In E. Bromet (Ed.), Long-Term Outcomes in Psychopathology Research: Rethinking the Scientific Agenda (pp. 140-157). New York, NY: Oxford University Press.

Kraemer, H. C., Kupfer, D. J., Clarke, D. E., Narrow, W. E., & Regier, D. A. (2012). DSM-5: how reliable is reliable enough?. *American Journal of Psychiatry, 169*(1), 13-15. https://doi.org/10.1176/appi.ajp.2011.11010050

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. https://doi.org/10.1002/wps.20566

Lundahl, B. W., Kunz, C., Brownell, C., Tollefson, D., & Burke, B. L. (2010). A meta-analysis of motivational interviewing: Twenty-five years of empirical studies. *Research on Social Work Practice, 20*(2), 137-160. https://doi.org/10.1177/1049731509347850

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. *Psychological Bulletin, 137*(5), 856. https://doi.org/10.1037/a0023678

Meehl, P. E. (1986). Diagnostic taxa as open concepts: Metatheoretical and statistical questions about reliability and construct validity in the grand strategy of nosological revision. In T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology* (pp.215–231). New York: Guilford Press.

Moore, T. V. (1930). The empirical determination of certain syndromes underlying praecox and manicdepressive psychoses. *The American Journal of Psychiatry*, 9, 719-738. http://dx.doi.org/10.1176/ajp.86.4.719

Morey, L. C. (2007). *Professional manual for the Personality Assessment Inventory (2nd ed.).* Odessa, FL: Psychological Assessment Resources.

Morey, L. C., Skodol, A. E., & Oldham, J. M. (2014). Clinician judgments of clinical utility: A comparison of DSM-IV-TR personality disorders and the alternative model for DSM-5 personality disorders. *Journal of Abnormal Psychology, 123*(2), 398-405. https://doi.org/10.1037/a0036481

Mullins-Sweatt, S. N., Lengel, G. J., & DeShong, H. L. (2016). The importance of considering clinical utility in the construction of a diagnostic manual. *Annual Review of Clinical Psychology, 12*, 133-155. https://doi.org/10.1146/annurev-clinpsy-021815-092954

Ormel, J., Raven, D., van Oort, F., Hartman, C. A., Reijneveld, S. A., Veenstra, R., ... & Oldehinkel, A. J. (2015). Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. *Psychological Medicine*, 45, 345-360. https://doi.org/10.1017/S0033291714001469

Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., ... & Fanous, A. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular Psychiatry, 21*(10), 1391-1399. https://doi.org/10.1038/mp.2015.197

Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 Field Trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *American Journal of Psychiatry, 170*(1), 59-70. https://doi.org/10.1176/appi.ajp.2012.12070999

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E., … Zimmerman, J. (2018, August 18). Integrating a Dimensional, Hierarchical Taxonomy of Psychopathology into Clinical Practice. https://doi.org/10.31234/osf.io/r2jt6

Teesson, M., Slade, T., & Mills, K. (2009). Comorbidity in Australia: findings of the 2007 national survey of mental health and wellbeing. *Australian and New Zealand Journal of Psychiatry, 43*(7), 606-614. https://doi.org/10.1080/00048670902970908

Wakefield, J.C. (2004). The myth of open concepts: Meehl's analysis of construct meaning versus black box essentialism. *Applied & Preventive Psychology*, 11, 77-82. https://10.1016/j.appsy.2004.02.014

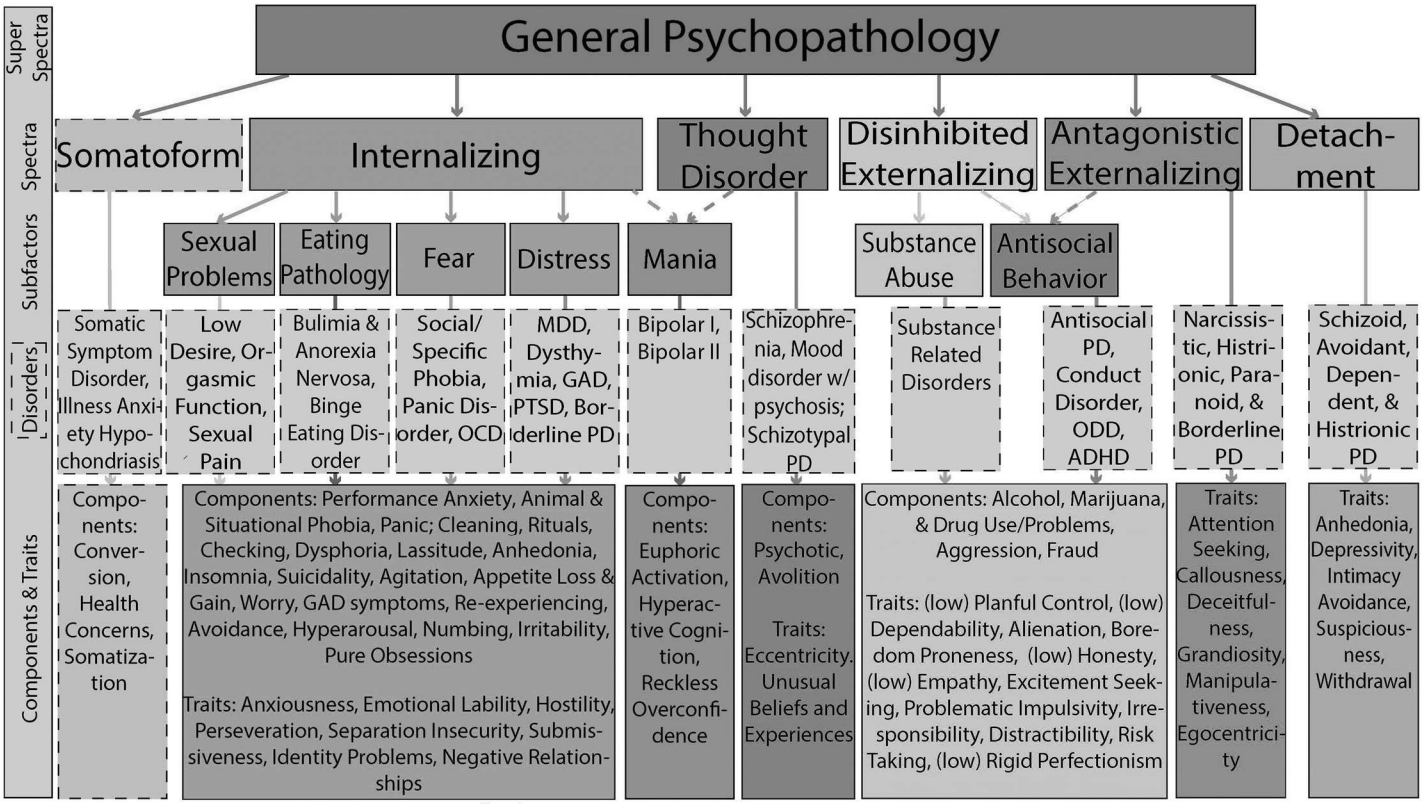
Waszczuk, M. A., Zimmerman, M., Ruggero, C., Li, K., MacNamara, A., Weinberg, A., ... & Kotov, R. (2017). What do clinicians treat: Diagnoses or symptoms? *Comprehensive Psychiatry, 79*, 80-88. https://doi.org/10.1016/j.comppsych.2017.04.004

Wright, A., Krueger, R. F., Hobbs, M. J., Markon, K. E., Eaton, N. R., & Slade, T. (2013). The structure of psychopathology: toward an expanded quantitative empirical model. *Journal of Abnormal Psychology, 122*(1), 281-294. https://doi.org/10.1037/a0030133

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry, 56*, 29-34. https://doi.org/10.1016/j.comppsych.2014.09.007

*Figure 1*

HiTOP model (reprinted with permission from Kotov et al., 2017)




*Note*. Not all disorders, components, and traits are represented in the figure.