

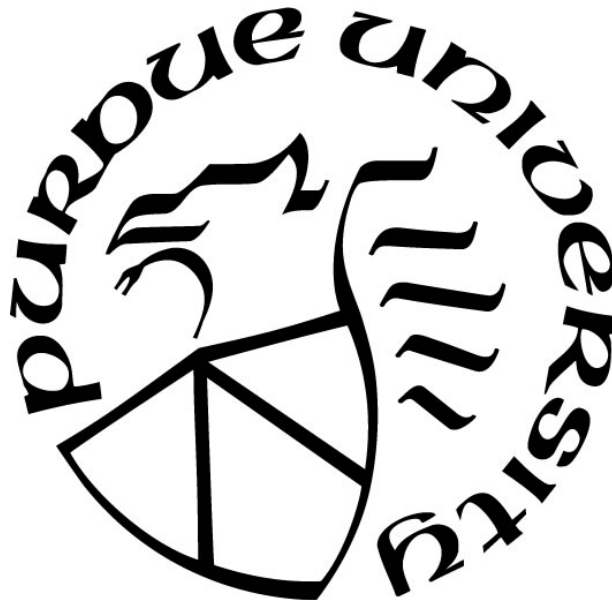
**CLINICIAN PERCEPTION OF THE CLINICAL UTILITY OF THE
HIERARCHICAL TAXONOMY OF PSYCHOPATHOLOGY (HiTOP)
SYSTEM**

by
Caroline E. Balling

A Thesis

*Submitted to the Faculty of Purdue University
In Partial Fulfillment of the Requirements for the degree of*

Master of Science



Department of Psychological Sciences
West Lafayette, Indiana
December 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Douglas B. Samuel, Chair

Department of Psychological Sciences

Dr. Susan C. South

Department of Psychological Sciences

Dr. Donald R. Lynam

Department of Psychological Sciences

Approved by:

Dr. Kimberly P. Kinzig

Dedicated to the recipients of mental health care and their providers.

TABLE OF CONTENTS

LIST OF TABLES	5
ABSTRACT.....	6
INTRODUCTION	7
METHODS	13
Participants.....	13
Procedure	14
Procedure for Applying Diagnostic Systems	14
Measures	15
Vignette Development	16
Data Analysis	17
Power Analysis	18
RESULTS	19
DISCUSSION	31
Limitations and Future Directions	33
Conclusion	35
REFERENCES	37
APPENDIX A.....	43
APPENDIX B	44
APPENDIX C	46
APPENDIX D.....	48
APPENDIX E	52

LIST OF TABLES

Table 1. Demographic Characteristics of Clinician Participants	20
Table 2. Professional Characteristics of Clinician Participants	21
Table 3. Comparing Clinical Utility Scores for DSM and HiTOP when Controlling for Vignette Assignment	23
Table 4. Correlations Between Clinician Factors and Overall CUQ Score	27
Table 5. Correlations Between Clinician Factors and all DSM CUQ Ratings	28
Table 6. Correlations Between Clinician Factors and all HiTOP CUQ Ratings	29

ABSTRACT

The standard of diagnosing and categorizing mental disorders in the United States has long been the Diagnostic and Statistical Manual of Mental Disorders (DSM), but the DSM has been criticized through evidence suggesting it lacks appropriate validity, reliability, and clinical utility. The Hierarchical Taxonomy of Psychopathology (HiTOP) has been offered as a solution to these criticisms. But the recommendation to replace the DSM and its categorical diagnostic system has been met with doubt and criticism by others in the field. A common sentiment in these critiques is a lack of evidence that the HiTOP dimensions are clinically useful or that clinicians would be open to applying them to their patients. The goal of the present study was to compare clinician perceptions of the HiTOP and DSM systems for the conceptualization of clinical cases. A sample of actively practicing clinicians ($n = 143$) rated one of three clinical vignettes using the HiTOP and DSM systems then rated the two approaches on seven indices of clinical utility. HiTOP was favored for overall clinical utility score as well as utility for formulating effective intervention, communicating clinical information to the client, comprehensively describing client psychopathology, describing global functioning, and ease of applying the system to the individual. There was no preference between HiTOP and the DSM for communicating with other mental health providers. The DSM was not favored for any clinical utility outcome. These results suggest interest in HiTOP and dissatisfaction with the DSM among clinicians.

INTRODUCTION

The standard of diagnosing and categorizing mental disorders in the United States is and has long been the Diagnostic and Statistical Manual of Mental Disorders (DSM), disseminated by the American Psychiatric Association (APA; APA, 2013). The manual has been translated into 24 languages and, in addition to its prominence in the United States, is increasingly the standard of diagnosis in Europe and Asia (APA Publishing, 2017; Kawa & Giordano, 2012). The influence of the DSM cuts across various professional disciplines; it is referred to by researchers, clinicians from multiple orientations, policymakers, criminal courts, and third-party reimbursement entities (APA Publishing, 2017; Kawa & Giordano, 2012).

A classification system such as the DSM accomplishes many goals. It is fundamental to the study and treatment of psychopathology because it provides a common language for researchers, clinicians, educators, and students to communicate with one other. It serves as a central tool for the dissemination of research, conceptualization of psychiatric cases, and education of future mental health workers. It permeates beyond the boundaries of the psychological and psychiatric disciplines and allows for interdisciplinary discourse, cooperation from insurance companies, and a foundation for understanding amongst the general public. In the United States, the DSM has been positioned to fill these roles. In this way, the influence and impact of the DSM is undisputable.

With this influence in mind, it is of concern that the DSM has been criticized through evidence suggesting it lacks appropriate validity, reliability, and clinical utility (Kotov et al., 2017; Widiger & Simonsen, 2005; Widiger et al., 2018). One might review the DSM-5 field trials as evidence of this: two of the most commonly occurring DSM diagnoses, generalized anxiety disorder and major depressive disorder, garnered unacceptably low kappa coefficients between 0.20 and 0.25 (Regier et al., 2013).

The limitations of the DSM approach to diagnosis are numerous. To start, most diagnoses are partly or fully polythetic, meaning that the diagnosis is made if a certain number of diagnostic criteria are endorsed from a longer list. Defining psychological disorders in this way produces significant heterogeneity within diagnostic groups due to the multiple permutations of symptom combinations that yield a positive diagnosis. This leads to the loss of valuable clinical information about a client, such as by eliminating details about what symptoms the client is experiencing, the

onset and chronicity of the individual's symptoms, and the severity of those indicators (Carragher et. al, 2015). As an example, an individual can meet DSM-5 criteria for borderline personality disorder (BPD) by endorsing five or more diagnostic criteria out of a possible nine criteria. Two clients might share the experience of only one of the nine BPD criteria and still be grouped under the same diagnostic label of BPD, despite what are undoubtedly distinct clinical symptom sets that could also vary in severity and course.

An additional consequence of the DSM system is that comorbidity is exceedingly common in clinical and community samples (Conway et al., 2019; Kotov et. al, 2017; Ruggero et al., 2019). This excessive co-occurrence of diagnoses challenges the notion that the DSM's categorical diagnoses are discrete entities. Importantly, this high comorbidity also complicates and obscures empirical and clinical work. For example, clinical trials that employ exclusion criteria barring potential participants with comorbid diagnoses can render these study samples largely ungeneralizable (e.g., Zimmerman et al., 2019).

Another concern with the DSM system is that most patients are categorized with ambiguous "unspecified" or "not otherwise specified" diagnoses because they do not meet the often-arbitrary diagnostic thresholds for the more specific disorders (Achenbach, 2015; Carragher et. al, 2015; Kotov et al., 2017). Of course, these unspecified diagnoses have minimal utility in research or clinical practice, as they provide minimal information about an individual's clinical presentation. This majority of clients carrying unspecified diagnoses calls into question the validity of the cutoffs that the DSM applies to its symptoms and syndromes.

Many criticisms of the DSM are rooted in the manual's assumption that mental disorders and associated symptoms are discrete categorical phenomena, such that they can be definitively identified as present or absent. This dichotomous format implies the existence of meaningful cutoffs for these constructs, but this has not been demonstrated empirically. On the contrary, evidence has repeatedly indicated that psychopathology exists on continua that include a full range of severity as well as adaptive and normative functioning (Achenbach, 2015; Carragher et. al, 2015; Eaton et al., 2011; Kotov et al., 2017). Taken together, this body of research has led some in the field to conclude that the DSM and its categorical diagnostic system are inadequate (Conway et al., 2019; Kotov et al., 2017; Widiger et al., 2018).

The Hierarchical Taxonomy of Psychopathology (HiTOP) has been offered as a solution to these criticisms (Kotov et al., 2017; Sauer-Zavala, in press; Widiger et al., 2018). HiTOP was

introduced in 2017 as a system for dimensionally classifying all variants of psychopathology, including Axis I and Axis II disorders (Kotov et al., 2017). Broadly, this model conceptualized the structure of psychopathology as a hierarchy: at the bottom of the hierarchy are psychiatric signs and symptoms, and at the top are increasingly broad dimensions such as internalizing and externalizing factors (Caspi et al., 2014; Kotov et al., 2017; Lahey et al., 2012). The highest and broadest level of the hierarchy is the superspectrum of general psychopathology (or *p* factor). At the next level down are the spectra such as internalizing and externalizing psychopathology, followed by subfactors such as fear and distress, and then the level of syndromes and disorders. At the bottom of the hierarchy are psychiatric signs, symptoms, components, and traits, such as worry, checking behaviors, and anhedonia. Importantly, HiTOP does not differ largely from the DSM in what are signs and symptoms of mental illness; instead, HiTOP represents an evidence-based reorganization of these lower-level constructs.

The debut publication of HiTOP in 2017 challenged many long-accepted, -researched, and -implemented ideologies in the fields of psychiatry and clinical psychology regarding the classification of psychopathology. Although this publication was a first encounter with dimensional theory for many professionals, the publication was built on a sturdy, long-standing foundation of research into the dimensional nature of psychopathology across the lifespan (Achenbach, 1966; Achenbach, 2015; Carragher et al., 2015; Krueger & Eaton, 2015). Indeed, the developmental psychopathology approach already operates with an understanding that psychopathology is predominantly dimensional in nature (e.g., Rutter, 2013; Rutter & Uher, 2012).

But the recommendation to replace the DSM and its categorical diagnostic system has been met with criticism and doubt by others in the field. One scholar likened the movement away from the DSM to “throwing the baby out with bath water” (Zimmerman, in press) and others stated that “the HiTOP consortium is writing checks it can’t cash” (Haefffel et al., in press). A common sentiment in these critiques is that there lacks evidence that the HiTOP dimensions are clinically useful or that clinicians will be able to apply them. Critics have also hypothesized resistance from clinicians to adopt such a system. These are certainly reasonable concerns.

Importantly, there is no data regarding clinician perception of the full HiTOP model. Data pertaining to the clinician perception of and satisfaction with HiTOP-aligned measures such as the ASEBA, MMPI, and PAI is also lacking, only furthering the uncertainty of how clinicians would perceive such an approach. There is a growing body of research regarding the clinician utility of

dimensional personality pathology models, which is one component of the HiTOP system. Past research has evidenced that clinicians find dimensional models of personality pathology acceptable and often preferred to the categorical conceptualization of the DSM (e.g., Bornstein & Natoli (2019); Glover et al., 2012; Hansen et al., 2019; Morey et al., 2014). For example, studies have found that clinicians consider dimensional personality traits to be more useful than categorical DSM diagnoses for clinical decision-making, treatment planning, comprehensively covering client difficulties, communication with the client, and generating global personality descriptions (Samuel & Widiger, 2006; Samuel & Widiger, 2011). One meta-analysis (Bornstein & Natoli, 2019) examined the clinical utility ratings of categorical and dimensional approaches to rating personality pathology across eleven studies, and the authors concluded that dimensional models were favored across the majority of indices of clinical utility. But even research into the clinical utility of dimensional personality models is a relatively new area of empirical inquiry, with the first publication on the topic debuting within the last two decades (Sprock, 2003).

It is noteworthy that the clinical utility of the DSM has not been adequately surveyed, either. Over a decade ago, Mullins-Sweatt and Widiger (2009) detailed this concern, noting, “an entirely valid DSM could actually have relatively little clinical utility if it is not feasible for usage in general clinical practice” (p. 310). Despite this limitation, most studies of the DSM focus on the reliability and validity of diagnoses while neglecting clinical utility (e.g., Clarke et al., 2013). One study, accomplished for the APA’s DSM-5 field trials, concluded that the DSM-5 was feasible and clinically useful (Mościcki et al., 2013). This was inferred from the findings that among clinicians of adult clients, 46% identified the DSM-5 as very or extremely easy for assessment, and 46% felt it was useful or extremely useful for routine clinical practice (Mościcki et al., 2013). Beyond this study, empirical information regarding clinician perception of the full DSM system is lacking.

Nonetheless, the questions and critiques surrounding the clinical application of HiTOP are certainly reasonable, and the clinician perception of this diagnostic approach is one area of sorely needed data. Indeed, in moments of differing opinions amongst researchers, perhaps it is most appropriate to consider the perspective of those most closely involved in patient care. The voices of clinicians are welcome and needed in the ongoing empirical shift away from the DSM, and we particularly need to assess the clinician perception of the utility, accessibility, and helpfulness of the full HiTOP model (Ruggero et al., 2019; Tyrer, 2018).

Accordingly, the goal of the present study was to assess the clinician perception of the HiTOP and DSM systems for the conceptualization of vignettes drawn from real clinical cases. Clinicians rated one of three clinical vignettes according to the three highest-order HiTOP levels: superspectra (i.e., the p factor), spectra (somatoform, internalizing, thought disorder, etc.), and subfactors (e.g., eating pathology, substance abuse, mania). These levels of the hierarchy were prioritized because these are the components of the system that differ most from the DSM, while the fourth and fifth levels contain constructs that DSM users would be familiar with. Accordingly, because there is limited distinction between the diagnoses of the DSM and the syndrome/disorder level of HiTOP at present, the fourth level of HiTOP was not assessed. The fifth level of HiTOP, symptom components and maladaptive traits (e.g., anxiousness, checking behaviors, avolition), was also not assessed. This was also in part because signs, symptoms, and components were already presented in the vignettes, so to have clinicians rate these would have been redundant. Additionally, the fifth level includes over eighty constructs, and that was not feasible for clinicians to rate on a brief survey with minimal compensation.

In addition to rating the client on the three levels of HiTOP, clinicians rated the client according to all possible DSM-5 diagnoses. Notably, whether the use of the DSM was typical for any given clinician, this approach largely dominates clinical training and billing for services, so all clinicians were likely familiar with the DSM approach and associated diagnoses. Following diagnostic ratings, the clinicians completed surveys assessing their subjective satisfaction with and overall perception of the HiTOP and DSM systems, as well as how the dimensional approach compared to their typical diagnostic categorization and their experience with the DSM.

Given the paucity of data comparing the DSM and HiTOP amongst practitioners, forming hypotheses proved challenging. Of course, the DSM carries the advantages of familiarity and existing infrastructure. Alternatively, because HiTOP is advertised as a solution to the DSM's shortcomings, it seemed feasible that the respondents would detect some of these benefits, such as the reconceptualization of the DSM's categorical comorbidity. Additionally, prior research investigating clinician perception of dimensional versus categorical personality models evidenced a clinician preference for dimensions (e.g., Bornstein & Natoli, 2019; Samuel & Widiger, 2006; Samuel & Widiger, 2011), which suggests the HiTOP approach might be preferred over the DSM in the present study. When considering the somewhat polarizing response to HiTOP amongst researchers, it also seemed possible that the clinicians might be similarly divided. With these

theoretical and empirical considerations in mind, we hypothesized that clinicians would rate the HiTOP model equally or more favorably than the DSM on measures of clinical utility.

METHODS

The study was preregistered on OSF prior to the collection of data. The preregistration can be accessed through the following link: <http://osf.io/n34wp>. There are not methodological deviations from the preregistration to report, beyond the addition of exploratory analyses that are identified as such in the manuscript. The study was approved by the affiliated university's Institutional Review Board (Protocol number: IRB-2021-713; Study Title: "Clinician Perception of the Utility of the Hierarchical Taxonomy of Psychopathology (HiTOP) System").

Participants

The study was advertised to clinicians via several electronic channels. This included emailing the clinical training directors of US Veterans Affairs (VAs), USA-based training directors of the Association of Psychology Postdoctoral and Internship Centers (APPIC), members of the Indiana Psychological Association (IPA), and members of the National Association for Training Clinic Directors (NATCD). The investigators also shared the study via their professionally affiliated Twitter accounts and on the email listserv and Facebook page for the Association for Behavioral and Cognitive Therapies (ABCT).

Two-hundred thirty-seven licensed and actively practicing clinicians entered the study. Of these, 63 participants dropped out of the survey without completing any outcome items. Thirty Facebook bot responses were eliminated after the authors observed identical and off-topic repeat-responses in the free-response variable of the survey (e.g., ten responses stating, "We can collect some questions about the mental illness of teenagers in the community;" six responses stating, "The psychological changes of some people in a specific environment"), all of which were responses to the Facebook group advertisement. An inspection of the data for these respondents suggested exceptionally inaccurate if not random responding on survey items, further justifying their deletion (e.g., selecting almost every DSM diagnosis including neurocognitive, sexual, psychotic, and other disorders for a given vignette). Finally, as was pre-registered, respondents who completed the survey in under four minutes were excluded for a lack of effortful responding, as it should take more than four minutes to read the vignette and rate the client on two diagnostic systems ($n = 1$). This amounted to a final sample size of 143 clinicians.

Procedure

After consenting to participate, respondents confirmed their status as active, practicing, mental health providers who had completed a graduate degree in a mental health-related field. If a respondent did not endorse these characteristics, the survey ended.

Qualifying clinicians read brief introductions to the two diagnostic systems (see Appendix A). Next, they read one of three randomly assigned clinical vignettes, with the vignette populating in a new window for continued reference throughout the study. The clinicians then made diagnostic ratings of the described client according to the HiTOP and DSM-5 approaches. Finally, the clinicians rated the clinical utility of both approaches. The order in which clinicians were asked to diagnose using each system was counterbalanced. After study completion and as compensation, respondents had the option to enter a raffle for one of thirty \$20 gift cards to their choice of Amazon, Etsy, or Target.

Procedure for Applying Diagnostic Systems

The HiTOP consortium is in the process of developing a comprehensive, omnibus HiTOP assessment, but it is not expected to debut for some time (Ruggero et al., 2019; Simms et al., in press). Even if such a measure were ready for use, it would likely be too long to reasonably expect clinicians to complete for the present study. Instead, the clinicians read brief descriptions of the various constructs then rated if, for the client depicted in the vignette, each construct was a) not at all a problem, b) somewhat of a problem, or c) very much a problem. This response format was modeled after the scoring system for the gold standard of diagnosing DSM psychopathology, the Semi-structured Clinical Interview for the DSM (SCID), in which raters score a clinical construct on a three-point range including absent, subthreshold, and threshold. Definitions and ratings were prepared for all components of the HiTOP model¹, including: superspectrum (general psychopathology/*p* factor), spectra (somatoform, internalizing, thought disorder, externalizing disinhibited, externalizing antagonistic, detachment), and subfactors (sexual problems, eating problems, fear, distress, mania, substance use, antisocial behavior; see Appendix B).

¹ For readers unfamiliar with HiTOP, it is recommended they inspect the visual aids displaying the hierarchy on p. 87 of Kotov et al. (2021) or p. 1072 of Ruggero et al. (2019).

The presentation of HiTOP constructs to rate began at the top of the hierarchy, with the superspectrum. The survey used skip logic such that the clinicians were only asked to score constructs that fell beneath the umbrella of difficulties the clinicians identified as “somewhat” or “very much” a problem on the previous level of the hierarchy. This followed the intended stepwise, cascading approach encouraged by the clinical workgroup of the HiTOP consortium: clinicians began at the higher levels, then worked downward. For example, if out of the five spectra, a clinician only indicated that the internalizing spectra was a problem for the client, then the clinician only rated the subfactors of the internalizing spectra: sexual problems, eating problems, fear, and distress. They were not asked to rate the mania, substance abuse, or antisocial behavior subfactors of the other spectra.

When the clinicians were asked to categorize the client according to the DSM-5, they were invited to use any reference materials (e.g., their DSM). The diagnostic options were organized to reflect the table of contents of the DSM-5, such that there were separate, labeled lists for each of the following diagnostic categories: neurodevelopmental disorders, schizophrenia and related disorders, bipolar and related disorders, depressive disorders, anxiety disorders, obsessive-compulsive and related disorders, trauma and stressor-related disorders, dissociative disorders, somatic symptom disorders, feeding and eating disorders, elimination disorders, sleep-wake disorders, sexual dysfunctions, gender dysphoria, impulse-control, and conduct disorders, neurocognitive disorders, substance-related and addictive disorders, personality disorders, paraphilic disorders, and other mental disorders (Appendix C, APA, 2013). The clinicians could select as many diagnoses as they desired, and they did not have to make a selection for categories they perceived as irrelevant.

Measures

The study’s primary outcome was the clinical utility of the target diagnostic approach as rated on the Clinical Utility Questionnaire (CUQ, First et al., 2004; Samuel & Widiger, 2006). Clinicians rated items on a 1- to 5-point Likert scale, where 1 indicated “not at all useful” and 5 indicated “extremely useful.” Clinicians completed this for both the HiTOP and DSM-5 systems. This measure includes six items that assess the utility of the target diagnostic system for: communicating information about the client with other mental health professionals, ease of applying the system to the individual, communicating with the client about themselves,

comprehensively describing the individual's psychopathology, formulating effective intervention for the client, and describing the individual's global functioning. Composite scores for the overall clinical utility of the HiTOP and the DSM approaches were generated by summing the CUQ item scores. The total score was out of a possible 30 points in which higher scores indicate greater utility.

Clinicians provided demographic information including their age, gender identity, and race or ethnicity. Additionally, a variety of professional characteristics of the clinicians were collected, such as theoretical orientation, typical methods of assessment, and years of clinical experience (Table 2). We also collected information regarding the clinicians' familiarity with the DSM and HiTOP systems and whether they had a preference between the two approaches.

Finally, at the end of the survey, the respondents had the option to provide any thoughts on the topics of the study in a free response box.

Vignette Development

The clinical vignettes were intended to reflect the types and amount of information that would be gleaned from a real-world therapy intake. Clinical information was to be presented in terms of client-specific thoughts, behaviors, and problems as clients share during an intake, rather than in terms of traditional diagnostic criteria. Further, vignettes sculpted to be prototypic examples of target DSM diagnoses were avoided because such prototypic, single-diagnosis cases are not reflective of the majority of outpatient clinical presentation (Zimmerman et al., 2007; Zimmerman, 2016). Instead, the investigators aimed towards more diagnostically complex cases than this. These objectives also aligned with the recommendation from Bornstein and Natoli's meta-analysis that future clinical utility research incorporate real clinical cases (Bornstein & Natoli, 2019).

Such vignettes were not readily available in the literature. Instead, the authors derived the vignettes from published clinical case studies to meet these goals and minimize investigator bias during vignette development. The principal investigator examined the table of contents of every issue of the *Clinical Case Studies* journal over the past five years (i.e., 2016-2021). Publications that suggested some level of case complexity instead of a single diagnosis (e.g., titles referencing transdiagnostic treatment or diagnostic comorbidity) were screened for consideration as a vignette template. The three case studies that provided the most detail about psychosocial history and psychiatric symptoms were selected as the bases for the vignettes: Lui, 2017; Scheiderer et al.,

2017; and Smith et al., 2020. Three vignettes were developed to ensure observed effects were not attributable to the characteristics of a certain case.

The authors consolidated these three case studies into one-page vignettes (see *Appendix D*) to minimize the time commitment of participants. The vignettes described the client's gender, age, occupation, presenting problems, current psychiatric symptoms, interpersonal and occupational functioning, and symptom chronicity and course. To maximize experimental control, age and gender were held constant across vignettes. Previous data indicates that among recipients of mental health care in the United States, the majority of clients are female and between the ages of 25 and 39, or on average about 32 years of age (SAMHSA Center for Behavioral Health Statistics and Quality, 2020). Thus, all clients were depicted in the vignettes as 32-year-old women. In describing psychiatric symptomatology, the investigators generally avoided traditional diagnostic labels (e.g., "Major Depressive Disorder") and grouped criteria lists (e.g., "feelings of emptiness, affective instability, marked impulsivity, chronic suicidality"). Across vignettes, certain details of symptomatology were expanded by the principal investigators based on information in the case study to more closely mirror how a client would present the information (e.g., altering "sleep difficulties" to "waking frequently throughout the night and not being able to go back to sleep, sometimes waking for the day as early as 3 or 4am").

Data Analysis

Analyses were conducted in SPSS Version 26 (IBM, 2019). For statistical inferences, alpha was set at .05, two-tailed. P-value criterion for exploratory analyses were adjusted using a False Discovery Rate (FDR) correction (Benjamini & Hochberg, 1995) to control for Type I error. Partial eta-squared effect sizes are reported for all main effects, where $\eta_p^2 \geq 0.01$ indicates a small effect, $\eta_p^2 \geq 0.06$ indicates a medium effect, and $\eta_p^2 \geq 0.14$ a large effect (Cohen, 1988). There was no need for treating missing items because there were no missing items on the CUQ across all responses. Possible outliers were retained in all analyses.

This was a mixed design with both within- and between-subjects elements. The primary analysis was a 2 (diagnostic model: DSM vs. HiTOP) x 3 (vignette assignment: A, B, or C) mixed model analysis of variance (ANOVA) with diagnostic model as the within-subjects variable, vignette assignment as the between-subjects variable, and clinical utility scores as the outcome. All pair-wise comparisons and simple effects were inspected.

Information from the questions pertaining to clinicians' demographic and professional characteristics were inspected descriptively. When applicable, these variables were correlated with the utility ratings for the DSM and HiTOP approaches.

Power Analysis

Prior to data collection, G*Power (Faul et al., 2007) was used to conduct power analyses. Our aim was to achieve 90% power to detect a small effect of Cohen's $f = 0.15$ (f values between .10 and .24 are considered "small") at the standard .05 alpha error probability. In our sample, this effect size was equivalent to a η_p^2 of approximately $\eta_p^2 = .03$. This target effect size was necessarily conservative given that the extant literature of clinician comparison of categorical versus dimensional personality models has yielded anywhere from negligible to large effects (e.g., Samuel & Widiger, 2006; Glover et al., 2012; Hansen et al., 2019; Morey et al., 2014). To detect a difference between the HiTOP and DSM clinical utility ratings using a 2 (diagnostic system) x 3 (vignette assignment) mixed model ANOVA where there is no correlation between the two measures, G*Power suggested a sample size of $n = 237$.

A sample size of 237 clinicians was the best-case scenario. With a less conservative target of 80% power to achieve a small effect of Cohen's $f = 0.2$, the target sample size was $n = 102$. Of note, this retained the specification that the repeated measures were not inter-correlated. If these values were intercorrelated even by .1, the total sample size specified by G*Power dropped to $n = 93$; a correlation of .3 dropped the sample to $n = 72$, and so on. In our sample, the repeated measures were correlated at $r = -.14$. A post-hoc sensitivity analysis revealed that with this correlation of $r = -.14$, the present study had 80% power to detect $f = 0.178$, a small effect. This suggests that we were adequately powered to detect small effects within the obtained sample.

RESULTS

Descriptive characteristics of the sample of clinicians are detailed in Table 1 and their professional characteristics are listed in Table 2. The sample was comprised of primarily female (62.7%) and White (86.7%) clinicians from the Clinical Psychology discipline (77.6%) with a mean age of 40.74 years ($SD = 10.80$). The most frequently endorsed theoretical orientations were second-wave CBT (64.3%), third-wave CBT (53.2%), and interpersonal (29.4%). Most respondents reported working with adult clients (83.2%), and the most commonly reported primary setting for clinical work was private practice (20.3%), followed by VAs (16.8%), then academic medical centers (15.4%). The sample was primarily comprised of clinicians with a PhD (51.4%) or PsyD (32.3%). On average, the clinicians reported having 12.59 years ($SD = 9.15$) of clinical experience and spending 47.1% ($SD = 25.36$) of their working hours on direct client contact. Across participants, 87 reported having access to a DSM that they did not use during the study, 33 reported having access to a DSM that they used while completing the study, and 22 indicated that they did not have access to a DSM during the study.

A 2 x 3 Mixed Model ANOVA was conducted with diagnostic model as the within-subjects variable and vignette assignment as the between-subjects variable. Results broadly evidenced higher clinical utility ratings of the HiTOP system compared to the DSM system (Table 3). When the six items of the CUQ were summed for an overall utility score with a minimum possible score of six points and a maximum possible score of thirty points, the HiTOP system ($M = 20.52$, $SD = 4.81$) was rated more positively than the DSM system ($M = 17.71$, $SD = 4.63$) at a statistically significant level, $F(1, 140) = 23.79$, $p < .001$, $\eta_p^2 = 0.15$.

Table 1. Demographic Characteristics of Clinician Participants

	<i>n</i>	%
Gender Identity (<i>n</i> = 1 missing)		
Female	89	62.68
Male	52	36.62
Non-Binary	1	0.70
Race/Ethnicity*		
White	124	86.71
Latinx or Hispanic	7	4.90
East Asian	5	3.50
Black or African American	3	2.10
Native American/American Indian	3	2.10
Native Hawaiian or other Pacific Islander	3	2.10
South Asian	1	0.70
Other	2	1.40
	<i>M</i>	<i>SD</i>
Age	40.74	10.80

Note. * = participants could make multiple selections.

Table 2. Professional Characteristics of Clinician Participants

	<i>n</i>	%
Psychology Subfield*		
Clinical Psychology	111	77.6
Counseling Psychology	19	13.3
Family, Couple, and Child Psychology	16	11.2
Forensic Psychology	13	9.1
Clinical Social Work	12	8.4
Other	8	5.6
Health Psychology	7	4.9
Educational Psychology	5	3.5
Theoretical Orientation*		
Second wave CBT	92	64.3
Third wave CBT (e.g. ACT, DBT)	76	53.1
Interpersonal	42	29.4
Humanistic (Person-centered, Gestalt, Existential)	36	25.2
Eclectic	32	22.4
First wave Behavioral Therapy	25	17.5
Psychodynamic	23	16.1
Other	5	3.5
Age Range of Clients*		
Adults	119	83.2
Adolescents/young adults	92	64.3
Children	49	34.3
Geriatric adults	46	32.2
Primary Setting of Clinical Work		
Private practice	29	20.4
Veterans Affairs	24	16.9
Academic medical center	22	15.5

Table 2 continued

Group practice	19	13.4
University clinic	18	12.7
Psychiatric Hospital	12	8.5
Community Mental Health Center	5	3.5
Correctional Facility/Prison System	4	2.8
Children's school system	1	0.7
Other	8	5.6
Highest Degree in a Mental Health Discipline		
PhD	73	51.4
PsyD	46	32.4
Master's	11	7.7
MD/PhD	6	4.2
MD	3	2.1
NP	2	1.4
	<i>M</i>	<i>SD</i>
Years Clinical Experience	12.59	9.15
% Working Hours of Direct Client Contact	47.08	25.36

Note. * = participants could make multiple selections.

Table 3. Comparing Clinical Utility Scores for DSM and HiTOP when Controlling for Vignette Assignment

Outcome	Diagnostic System		Within-Subjects:			Between-Subjects:			Interaction:		
	<u>Diagnostic System</u>		<u>Diagnostic System</u>			<u>Vignette Assignment</u>			<u>Vignette Assignment</u>		
	DSM	HiTOP	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2	<i>F</i>	<i>p</i>	η_p^2
CUQ Total Score	17.71 (4.63)	20.52 (4.81)	23.79	< .001	0.15	0.55	0.578	0.01	3.74	0.026	0.05
Ease of applying system to individual	3.07 (.96)	3.36 (.95)	6.72	0.011	0.05	0.24	0.789	0.00	6.18	0.003	0.08
Communicating with other providers	3.36 (.98)	3.25 (1.01)	0.73	0.395	0.01	0.56	0.553	0.01	0.90	0.409	0.01
Communicating to individual	2.90 (.90)	3.40 (1.01)	17.31	< .001	0.11	0.56	0.564	0.01	1.83	0.164	0.03
Comprehensively describing psychopathology	2.77 (.94)	3.50 (1.01)	38.92	< .001	0.22	0.12	0.883	0.00	3.59	0.030	0.05
Formulating effective intervention	2.93 (.97)	3.47 (1.01)	21.38	< .001	0.13	0.99	0.374	0.01	5.91	0.003	0.08
Describing global functioning	2.69 (.95)	3.55 (.99)	49.71	< .001	0.26	0.51	0.601	0.01	0.68	0.510	0.01

We repeated the 2 x 3 ANOVA for each of the six individual CUQ items, as they assess incrementally informative components of utility. These analyses were listed as potential exploratory analyses in the preregistration, so p -values were adjusted using FDR correction. For each item, the minimum possible score was one point and the maximum was five points. The HiTOP system was rated more favorably for five out of six of the components of utility, with two effect sizes considered to be large, two medium, and one small. HiTOP scored higher than the DSM for describing global functioning [*Mean difference* = 0.86, $F(1, 140) = 49.71$, $p < .001$, $\eta_p^2 = 0.26$], comprehensively describing psychopathology [*Mean difference* = 0.73, $F(1, 140) = 38.92$, $p < .001$, $\eta_p^2 = 0.22$], formulating effective intervention [*Mean difference* = 0.54, $F(1, 140) = 21.38$, $p < .001$, $\eta_p^2 = 0.13$], communicating clinical information to the client [*Mean difference* = 0.50, $F(1, 140) = 17.31$, $p < .001$, $\eta_p^2 = 0.11$], and ease of applying the system to the individual [*Mean difference* = 0.29, $F(1, 140) = 6.72$, $p = .011$, $\eta_p^2 = 0.05$]. There was no statistically significant difference between the DSM and HiTOP systems for communicating information about the individual to other mental health providers [*Mean difference* = -0.11, $F(1, 140) = 0.73$, $p = .395$, $\eta_p^2 = 0.01$].

There were not statistically significant main effects of vignette assignment for any outcome, including: overall CUQ score [$F(2, 140) = 0.55$, $p = 0.578$, $\eta_p^2 = 0.01$], communicating information about the client with other mental health professionals [$F(2, 140) = 0.56$, $p = 0.553$, $\eta_p^2 = 0.01$], ease of applying the system to the individual [$F(2, 140) = 0.24$, $p = 0.789$, $\eta_p^2 = 0.00$], communicating with the client about themselves [$F(2, 140) = 0.56$, $p = 0.564$, $\eta_p^2 = 0.01$], comprehensively describing the individual's psychopathology [$F(2, 140) = 0.12$, $p = 0.883$, $\eta_p^2 = 0.00$], formulating effective intervention for the client [$F(2, 140) = 0.99$, $p = 0.374$, $\eta_p^2 = 0.01$], and describing the individual's global functioning [$F(2, 140) = 0.51$, $p = 0.601$, $\eta_p^2 = 0.01$].

There were statistically significant interaction effects between diagnostic system and vignette assignment for four of these seven analyses. All interaction and simple effects pertaining the analysis of individual CUQ items were adjusted using FDR correction. There was an interaction effect related to the overall CUQ score [$F(2, 140) = 3.74$, $p = .026$, $\eta_p^2 = 0.05$], such that there was a statistically significant difference between diagnostic systems for Vignette A (*Mean difference* = -5.13, $p < .001$), but not Vignette B (*Mean difference* = -1.76, $p = .081$), nor Vignette C (*Mean difference* = -1.67, $p = .101$). Regarding the interaction effect for the comprehensively describing psychopathology item [$F(2, 140) = 3.59$, $p = .030$, $\eta_p^2 = 0.05$], there

were still statistically significant effects between diagnostic systems for all vignettes: Vignette A (*Mean difference* = -1.17, $p < .001$), Vignette B (*Mean difference* = -0.43, $p = .035$), and Vignette C (*Mean difference* = -0.60, $p = .004$). There was also an interaction effect for ease of applying the system to the individual [$F(2, 140) = 6.18$, $p = .003$, $\eta_p^2 = 0.08$], and simple effect comparisons displayed a significant difference between diagnostic systems within Vignette A (*Mean difference* = -0.89, $p < .001$), but not Vignette B (*Mean difference* = 0.02, $p = .919$), nor Vignette C (*Mean difference* = -0.04, $p = .837$). Similarly, there was an interaction effect for formulating effective intervention outcome [$F(2, 140) = 5.91$, $p = .003$, $\eta_p^2 = 0.08$], and simple effect comparisons displayed a significant difference between diagnostic systems within Vignette A (*Mean difference* = -1.13, $p < .001$), but not Vignette B (*Mean difference* = -0.25, $p = .229$), nor Vignette C (*Mean difference* = -0.27, $p = .188$). There were not statistically significant interaction effects for describing global functioning [$F(2, 140) = 3.74$, $p = .026$, $\eta_p^2 = 0.05$], communicating to the individual [$F(2, 140) = 1.83$, $p = .164$, $\eta_p^2 = 0.03$], nor communicating with other providers [$F(2, 140) = 0.90$, $p = .409$, $\eta_p^2 = 0.01$].

The order in which participants rated the diagnostic systems was counter-balanced, and there were no observed ordering effects for ratings of the DSM, $t(141) = -.075$, $p = .940$, nor HiTOP, $t(141) = .15$, $p = .879$.

As was pre-registered, quantitative clinician factors including familiarity with the diagnostic systems, clinician age, years of clinical experience, and percent working hours spent in client contact were correlated with the CUQ outcomes (Tables 4, 5, and 6). Years of clinical experience displayed a positive relationship with familiarity with the DSM ($r = .27$, $p < .01$), and a negative relationship with familiarity with HiTOP ($r = -.30$, $p < .01$). Years of clinical experience revealed several statistically significant, negative relations with DSM outcomes: overall CUQ score ($r = -.22$, $p < .01$), communicating to the individual ($r = -.20$, $p < .05$), comprehensively describing psychopathology ($r = -.18$, $p < .05$), formulating effective intervention ($r = -.22$, $p < .01$), and describing global functioning ($r = -.23$, $p < .01$). There were also negative associations between clinician age and reported usefulness of the DSM for communicating to the patient ($r = -.17$, $p < .05$), as well as between clinician age and utility of the DSM for formulating effective intervention ($r = -.17$, $p < .05$). There was a positive association between familiarity with the DSM and reported ease of applying the DSM system ($r = .18$, $p < .05$). Only one statistically significant correlation between clinician factors and HiTOP-specific outcomes was observed, such that

familiarity with HiTOP was positively associated with reported ease of HiTOP for communicating to other mental health providers about the individual ($r = .18$, $p < .05$).

Table 4. Correlations Between Clinician Factors and Overall CUQ Score

	1	2	3	4	5	6	7
1. DSM CUQ Sum	—						
2. HiTOP CUQ Sum	-0.137	—					
3. Familiarity with DSM	0.08	0.14	—				
4. Familiarity with HiTOP	-0.11	0.13	-0.30**	—			
5. Clinician Age	-0.16	-0.06	0.09	-0.06	—		
6. Years of Clinical Experience	-0.22**	-0.03	0.27**	-0.09	0.88**	—	
7. % Working Hours of Direct Client Contact	0.11	-0.09	-0.05	-0.08	0.00	-0.07	—

Note. *significant at $p < .01$.

Table 5. Correlations Between Clinician Factors and all DSM CUQ Ratings

	1	2	3	4	5	6	7	8	9	10	11
1. DSM CUQ Sum	—										
2. Ease of Applying System to Individual	0.87**	—									
3. Communicating with Other Providers	0.83**	0.75**	—								
4. Communicating to Individual	0.79**	0.61**	0.60**	—							
5. Comprehensively Describing Psychopathology	0.84**	0.67**	0.61**	0.61**	—						
6. Formulating Effective Intervention	0.83**	0.70**	0.59**	0.58**	0.68**	—					
7. Describing Global Functioning	0.71**	0.49**	0.47**	0.46**	0.54**	0.48**	—				
8. Familiarity with DSM	0.08	0.18*	0.16	0.13	-0.01	0.07	-0.16	—			
9. Familiarity with HiTOP	-0.11	-0.15	-0.21*	-0.11	-0.07	-0.15	0.16	-0.30**	—		
10. Clinician Age	-0.16	-0.11	-0.12	-0.17*	-0.09	-0.17*	-0.13	0.09	-0.06	—	
11. Years of Clinical Experience	-0.22**	-0.13	-0.12	-0.20*	-0.18*	-0.22**	-0.23**	0.27**	-0.09	0.88**	—
12. % Working Hours of Direct Client Contact	0.11	0.05	0.07	0.14	0.08	0.07	0.12	-0.05	-0.08	0.00	-0.07

Note. *significant at $p < .05$, **significant at $p < .01$.

Table 6. Correlations Between Clinician Factors and all HiTOP CUQ Ratings

	1	2	3	4	5	6	7	8	9	10	11
1. DSM CUQ Sum	—										
2. Ease of Applying System to Individual	0.84**	—									
3. Communicating with Other Providers	0.81**	0.71**	—								
4. Communicating to Individual	0.76**	0.49**	0.53**	—							
5. Comprehensively Describing Psychopathology	0.86**	0.72**	0.64**	0.52**	—						
6. Formulating Effective Intervention	0.83**	0.67**	0.60**	0.59**	0.71**	—					
7. Describing Global Functioning	0.72**	0.48**	0.45**	0.54**	0.56**	0.44**	—				
8. Familiarity with DSM	0.14	0.16	0.06	0.04	0.16	0.17*	0.11	—			
9. Familiarity with HiTOP	0.13	0.11	0.18*	0.13	0.03	0.05	0.14	-0.30**	—		
10. Clinician Age	-0.06	-0.06	-0.03	-0.05	-0.04	-0.03	-0.10	0.09	-0.06	—	
11. Years of Clinical Experience	-0.03	-0.01	-0.03	-0.06	-0.01	-0.02	-0.03	0.27**	-0.09	0.88**	—
12. % Working Hours of Direct Client Contact	-0.09	0.00	-0.07	-0.01	-0.13	-0.04	-0.16	-0.05	-0.08	0.00	-0.07

Note. *significant at $p < .05$, **significant at $p < .01$.

Clinicians were also asked to rate how familiar they were with both diagnostic models, and 77 participants indicated that they were “not at all familiar” with HiTOP prior to participating in study. As an exploratory analysis aimed to assess potential response bias toward those familiar with HiTOP in our sample, we re-ran the 2 x 3 Mixed Model ANOVAs with only these 77 participants who had no prior exposure to HiTOP. All p -values were adjusted using FDR correction. Broadly, the direction of effects held across all seven analyses, and the patterns of statistical significance held across six. HiTOP was still preferred for the outcomes of overall CUQ score (*Mean difference* = 2.54, p = .001, η_p^2 = 0.13), describing global functioning (*Mean difference* = 0.97, p < .001, η_p^2 = 0.33), comprehensively describing psychopathology (*Mean difference* = 0.73, p < .001, η_p^2 = 0.24), formulating intervention (*Mean difference* = 0.48, p = .002, η_p^2 = 0.12), and communication with the client (*Mean difference* = 0.42, p = .017, η_p^2 = 0.08). As with the initial analyses, there was no effect of diagnostic system for communicating with other providers (*Mean difference* = -0.26, p = .094, η_p^2 = 0.04). The only change was that HiTOP (M = 3.34, SD = 1.00) was no longer favored over the DSM (M = 3.14, SD = 0.90) for ease of applying the diagnostic system to the individual (p = .218, η_p^2 = 0.02). As in the primary analyses described above, in this sub-sample, the DSM was not favored for any outcome.

When asked explicitly which model they would prefer for diagnosis in their clinical work, 73 clinicians (51.41%) indicated HiTOP, 54 (38.03%) the DSM, and 15 (10.56%) neither approach. Comments left by the clinicians in the free-response box at the end of the survey are presented in Appendix E and organized into comments about the DSM, about HiTOP, comparing both systems, and regarding the study itself. As examples, one clinician noted, “I love the idea of a dimensional approach to dx, but I don't know how to use HiTop. I think having a specific diagnosis is often helpful for patients, as it helps them understand that they are not alone and that there is a name for what they are experiencing.” Another stated, “I appreciate these alternate diagnostic models and find them more useful than the DSM, whose only value for me is getting the proper code for insurance and flattening out wrinkly papers.”

DISCUSSION

In the time since the debut of the HiTOP system, the clinical utility of HiTOP has been a crucial question of interest. Prior to the present study, there was no data to speak to this question. The present findings convey great potential for the clinical utility of the HiTOP system. As previous research has displayed clinician preference towards HiTOP-friendly approaches to personality pathology over the DSM's categorical PD diagnoses (e.g., Bornstein & Natoli, 2019; Hansen et al., 2019; Samuel & Widiger, 2011), the results of the present study reveal similar trends for the full diagnostic systems.

Among this sample of practicing clinicians, HiTOP was favored for its utility in formulating effective intervention, communicating clinical information to the client, comprehensively describing psychopathology, describing global functioning, and ease of applying the system to the individual. It also received a higher overall CUQ score. These findings suggest that clinicians perceived HiTOP as the better option for individualizing services, as well as a more accessible format for presenting diagnostic feedback to clients. The findings could also indicate that HiTOP appears the better option for capturing the complexity of a client's clinical presentation while offering an effective estimate of overall impairment, likely a reflection of the specificity of the lower hierarchy levels and the *p* factor at the highest hierarchy level. Indeed, these findings coalesce with previous postulations regarding the potential benefits of HiTOP in clinical settings (e.g., Sauer-Zavala, in press; Stanton et al., 2019).

The clinicians also found HiTOP to be less difficult to use than the DSM, perhaps suggesting that a diagnostic approach more reflective of HiTOP might increase engagement in the diagnostic process. Improving clinician participation in the diagnostic process is a notable area of growth in the mental health field (Ruggero et al., 2019). One 2014 survey found that in a sample of over 6,000 USA healthcare workers, 55% denied incorporating their DSM into their practice, including 41% of surveyed psychiatrists and 40% of surveyed psychologists (Cassels, 2017). Of course, a lack of time or incentive to participate in the diagnostic process might be an issue for any diagnostic modality (Ruggero et al., 2019), so the possibility of HiTOP improving engagement should be examined explicitly in the future.

The clinicians indicated no preference between HiTOP and the DSM for communicating information about the client to other mental health providers. It was somewhat unexpected that the

DSM was not preferred for this metric of utility. Of the six clinical utility constructs assessed with the CUQ, the DSM arguably had its best chance to out-perform HiTOP on this item because the DSM is the framework through which mental health care in the United States operates. Attempting to communicate with members of a DSM-entrenched system using a HiTOP lens could reasonably seem a difficult task.

This lack of preference for the DSM in communicating with other mental health providers could prompt concern that the sample was driven by a response bias towards those already familiar with HiTOP (and presumably favorable towards it). To investigate this possibility, we conducted exploratory analyses re-running the CUQ analyses and including only the sub-group of 77 clinicians who were completely unfamiliar with HiTOP prior to the study. The only observed change across the seven effects was that this subsample displayed no preference between systems for ease of applying the diagnostic system to the individual. HiTOP was preferred for five utility constructs, and the DSM was preferred for none. Thus, concern that our sample was biased towards HiTOP is unsupported.

Instead, it seems feasible that the clinicians understood that to use a HiTOP lens for inter-professional communication would not mark a dramatic change in the way they communicate. Indeed, HiTOP includes the signs, traits, symptoms, and components that make up many of the diagnoses listed in DSM, and the two systems represent different ways of organizing these constructs. A clinician might already communicate at the level of signs and symptoms, for example describing a client as experiencing ruminative worry, insomnia, and difficulty concentrating. This description is not DSM-specific. The lack of preference between HiTOP and the DSM for communicating with other providers might convey that clinicians do not rely on the DSM to communicate even with each other.

It is worth emphasizing the findings from the sub-group of HiTOP-naïve clinicians beyond the lens of the professional communication item. These findings are quite striking as they indicate that HiTOP is compelling and simple enough to be preferred over the DSM even amongst clinicians who were seeing it for the first time. This not only speaks optimistically of the promise of HiTOP, but also counters concerns about its overwhelming complexity (Haefel et al., in press). These results accentuate that clinicians readily grasp the HiTOP system and apply it to novel patients.

Although the primary goal of the present study was assessing the perceived clinical utility of HiTOP, the implications regarding the perceived utility of the DSM are also striking. The DSM was not favored for any clinical utility outcome across all analyses. This is of significant concern given the influence and impact of the DSM, and these results should raise significant alarm to those dedicated to the continued implementation of the DSM system. The standards must be higher for what remains as the standard of diagnosis across many disciplines. We hope that future DSM workgroups will consider this data in future iterations of the manual.

It is difficult to determine whether these quantitative HiTOP successes were driven by an appreciation of HiTOP, a dislike of the DSM, or some combination of both. Trends across the limited free-response information collected in this study do suggest a fairly equal combination of the two forces. In the future, structured qualitative surveys could be well-suited to arbitrate such comparisons.

Despite robust effects favoring HiTOP across the CUQ, when asked which diagnostic system clinicians would prefer for use in clinical practice, only the slightest majority favored HiTOP: 73 (51.41%) said HiTOP, 54 (38.03%) said the DSM, and 15 (10.56%) said neither. This reflects the substantial progress that must be made for either diagnostic system to succeed in the long-term.

Limitations and Future Directions

Vignette A seemed to carry the main effect of HiTOP preference for overall CUQ score, ease of applying the system to the individual, and formulating effective intervention, as clinicians assigned to read Vignettes B or C did not indicate a preference between HiTOP and the DSM. We did not observe interaction effects between diagnostic system and vignette assignment for comprehensively describing psychopathology, describing global functioning, communicating to the individual, nor communicating with other providers. Notably, simple effects comparisons never revealed a preference for the DSM; what varied was whether HiTOP was preferred or if there was no preference between HiTOP and the DSM. We did not forge systematic differences between the vignettes on purpose, so it is difficult to know why Vignette A was associated with certain higher utility scores for HiTOP. This vignette did require almost no elaborations or additions on behalf of the principal investigators, as its associated case study (Scheiderer et al., 2017) provided ample detail about the symptoms the client was experiencing. Perhaps the revisions

to vignettes B and C were biased towards the DSM on the part of the principal investigators given their training backgrounds. Or perhaps Vignette A came across as a less prototypic diagnostic case study. There is no doubt that prototypicality likely affects perceived utility of systems (e.g., Samuel & Widiger, 2011), but there remains a question of degree. Future research into meaningfully varied vignettes and associated effect differences would be a welcome area of exploration.

As another limitation of the study, the approaches for applying the two diagnostic systems were imperfect proxies. For example, engaging in the HiTOP approach was likely more time-consuming than with the DSM because the HiTOP approach incorporated introductory descriptions of each HiTOP construct, and this could have been of some detriment to the perception of HiTOP. Nonetheless, the diagnostic proxies accomplished the main goal of obligating the clinician to effortfully consider the DSM and HiTOP systems in relation to the client before making clinical utility ratings.

This area of research would benefit from a broader measure of clinical utility with interpretation guidelines. Though use of the CUQ has been a standard for diagnostic clinical utility research, the measure could be expanded to include domains such as usefulness for assessing risk or developing therapeutic alliance. Further, there are not cutoffs for interpreting CUQ scores or suggesting a minimal acceptable level of utility. Overall, HiTOP received an average CUQ score of 20.52 ($SD = 4.81$), and the DSM earned an average score of 17.71 ($SD = 4.63$). It is difficult to put this approximately three-point difference into context in terms of clinical impact. It is possible that out of a possible score of 30, both systems are underperforming, but more data is needed to make an interpretation such as this. Future research should explore cut-points that might reflect the strength of a system's clinical utility.

About three-quarters of our sample was comprised of mental health professionals from Clinical Psychology backgrounds. Future examinations of the clinical utility of HiTOP would benefit from obtaining ratings by other types of providers, particularly amongst psychiatrists. It is reasonable to suspect that psychiatrists would differ in their preferences for the DSM over HiTOP, particularly because the DSM is a product of the primary governing body of psychiatry in the United States, the APA. It would also be of interest to assess utility beyond providers specific to mental health care, such as primary care physicians, who often have a role in mental health treatment and associated referrals (Bornstein & Natoli, 2019). Further, this research must

eventually extend beyond vignettes and into psychiatric clinics and hospitals, to be applied to the actual clients of surveyed clinicians.

The ultimate measure of clinical utility will be to assess if a HiTOP approach improves treatment outcomes, and commentaries and criticisms of the system have made this clear (e.g., Tyrer, 2018; Zimmerman, in press). In the area of personality functioning, one study found that dimensionally-rated, HiTOP-friendly personality scores predicted aspects of psychological functioning after six months of substance use disorder treatment (Samuel & Widiger, 2011). Further exploration into the predictive validity of diagnostic models is of the utmost importance for future research, both for HiTOP and the DSM, as data is notably lacking in this area for both systems. It is worth emphasizing that the goal is to employ *whatever* system is most useful for improving the mental health of our clients. These are not merely academic questions.

Conclusion

Given the study's limitations and in the absence of other data, our response to these results is certainly not to suggest that HiTOP must urgently replace the DSM. In the meantime, for critics that assert HiTOP is too complicated, cumbersome, or unapplicable for clinicians, these data suggest this is unfounded. There are certainly reasonable concerns, raised in the absence of data. Now that at least this data exists, we hope that a portion of these apprehensions are alleviated. From this vantage point, it appears that when it comes to HiTOP, clinicians are capable and interested.

For the interested clinician, there are several manners of implementing the HiTOP approach as the field awaits formal assessment measures and associated guidelines. To read more about HiTOP and access a variety of clinical tools for its use (including billing) in practice, visit the HiTOP Clinical Network page at <https://hitop.unt.edu/introduction> (Ruggero, 2021). This also includes recommendations for HiTOP-friendly measures to use, such as the Personality Assessment Inventory PAI (PAI) or the Minnesota Multiphasic Personality Inventory (MMPI), as clinical field trials for a formal HiTOP assessment continue (Ruggero, 2021).

A 2018 *World Psychiatry* commentary concluded, “if [HiTOP] and colleagues can come forward with more clinical meat to add to their helping of science, things will certainly change” (Tyrer, 2018, p. 296). Ultimately, this single study does not serve as the definitive answer regarding the clinical utility of HiTOP or the DSM. However, the conclusion of the present study is that

actively practicing clinicians displayed robust preference for HiTOP over the DSM in measures of clinical utility. We hope this study is an encouraging serving.

REFERENCES

- Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied*, 80, 1-37. doi: 10.1037/h0093906
- Achenbach, T. M. (2015). Transdiagnostic heterogeneity, hierarchical dimensional models, and societal, cultural, and individual differences in the developmental understanding of psychopathology. *European Child & Adolescent Psychiatry*, 24, 1419-1422. doi: 10.1007/s00787-015-0795-0
- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA.
- American Psychiatric Association (APA) Publishing. (2017, April 17). *DSM-5 Collection Translation Publishers*. https://www.appi.org/getattachment/01bd83bf-bea4-4478-be13-fdb280af3539/APP_DSM-5_Translation_Publishers_Contact_Info.pdf
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bornstein, R. F., & Natoli, A. P. (2019). Clinical utility of categorical and dimensional perspectives on personality pathology: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, 10, 479-490. doi: 10.1037/per0000365
- Carragher, N., Krueger, R. F., Eaton, N. R., & Slade, T. (2015). Disorders without borders: Current and future directions in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology*, 50, 339-350. doi: 10.1007/s00127-014-1004-z
- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & 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>
- Cassels, C. (2017, August 16). *One third of psychiatrists not using DSM-5*. Medscape. Retrieved October 18, 2021, from <https://www.medscape.com/viewarticle/830099>

- Clarke, D. E., Narrow, W. E., Regier, D. A., Kuramoto, S. J., Kupfer, D. J., Kuhl, E. A., Greiner, L., & Kraemer, H. C. (2013). DSM-5 field trials in the United States and Canada, Part I: Study design, sampling strategy, implementation, and analytic approaches. *The American Journal of Psychiatry*, 170, 43-58. <http://dx.doi.org/10.1176/appi.ajp.2012.12070998>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.
- Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., Mullins-Sweatt, S. N., Shackman, A. J., Skodol, A. E., South, S. C., Sunderland, M., Waszczuk, M. A., Zald, D. H., Afzali, M. H., Bornovalova, M. A., Carragher, N., Docherty, A. R., Jonas, K. G., Krueger, R. F., . . . Eaton, N. R. (2019). A hierarchical taxonomy of psychopathology can transform mental health research. *Perspectives on Psychological Science*, 14(3), 419-436. <https://doi.org/10.1177/1745691618810696>
- Eaton, N. R., Krueger, R. F., South, S. C., Simms, L. J., & Clark, L. A. (2011). Contrasting prototypes and dimensions in the classification of personality pathology: Evidence that dimensions, but not prototypes, are robust. *Psychological Medicine*, 41, 1151-1163.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- First, M. B., Pincus, H. A., Levine, J. B., Williams, J. B. W., Ustun, B., & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnoses. *American Journal of Psychiatry*, 161, 946-954.
- Glover, N. G., Crego, C., & Widiger, T. A. (2012). The clinical utility of the Five Factor Model of personality disorder. *Personality Disorders*, 3, 176-184. <http://dx.doi.org/10.1037/a0024030>
- Haefel, G. J., Jeronimus, B. F., Kaiser, B. N., Weaver, L. J., Soyster, P. D., Fisher, A. J., Vargas, I., Goodson, J. T., & Lu, W. (in press). Folk classification and factor rotations: Whales, sharks, and the problems with HiTOP. *Clinical Psychological Science*.
- Hansen, S. J., Christensen, S., Kongerslev, M. T., First, M. B., Widiger, T. A., Simonsen, E., & Bach, B. (2019). Mental health professionals' perceived clinical utility of the ICD-10 vs. ICD-11 classification of personality disorders. *Personality and Mental Health*, 13, 84-95. [http:// dx.doi.org/10.1002/pmh.1442](http://dx.doi.org/10.1002/pmh.1442)

- IBM Corp. Released 2019. IBM SPSS Statistics for Macintosh, Version 26.0. Armonk, NY: IBM Corp
- Kawa, S., & Giordano, J. (2012). A brief historicity of the *Diagnostic and Statistical Manual of Mental Disorders*: Issues and implications for the future of psychiatric canon and practice. *Philosophy, Ethics, and Humanities in Medicine*, 7, 2. <https://doi.org/10.1186/1747-5341-7-2>
- Kotov, R., Krueger, R., Watson, D., Achenbach, T., Althoff, R., Bagby, M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K. . . . Zimmerman, M. L. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126, 454-477. <http://dx.doi.org/10.1037/abn0000258>
- Krueger, R. F., & Eaton, N. R. (2015). Transdiagnostic factors of mental disorders. *World Psychiatry*, 14, 27-29. doi: 10.1002/wps.20175
- 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. doi: 10.1037/a0028355
- Lui, P. P. (2017). Incorporating meta-emotions in integrative cognitive-affective therapy to treat comorbid bulimia nervosa and substance use disorders in a Latina American. *Clinical Case Studies*, 16(4), 328-345. <https://doi.org/10.1177/1534650117698038>.
- 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, 398-405. <http://dx.doi.org/10.1037/a0036481>
- Mościcki, E. K., Clarke, D. E., Kuramoto, S. J., Kraemer, H. C., Narrow, W. E., Kupfer, D. J., & Regier, D. A. (2013). Testing DSM-5 in routine clinical practice settings: Feasibility and clinical utility. *Psychiatric Services*, 64(10), 952-960. <https://doi.org/10.1176/appi.ps.201300098>
- Mullins-Sweatt, S. N., & Widiger, T. A. (2009). Clinical utility and DSM-V. *Psychological Assessment*, 21(3), 302-312. <https://doi.org/10.1037/a0016607>

- 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. *The American Journal of Psychiatry*, 170, 59-70. <http://dx.doi.org/10.1176/appi.ajp.2012.12070999>
- Ruggero, C. (2021). *HiTOP Clinical Network* . Introduction. Retrieved October 18, 2021, from <https://hitop.unt.edu/introduction>.
- Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E., Mullins-Sweatt, S. N., Patrick, C. J., Bach, B., Cicero, D. C., Docherty, A., Simms, L. J., Bagby, R. M., Krueger, R. F., Callahan, J. L., Chmielewski, M., Conway, C. C., De Clercq, B., Dornbach-Bender, A., . . . Zimmermann, J. (2019). Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. *Journal of Consulting and Clinical Psychology*, 87, 1069-1084. <https://doi.org/10.1037/ccp0000452>
- Rutter M. (2013). Developmental psychopathology: a paradigm shift or just a relabeling? *Development and Psychopathology*, 25, 1201-1213. <https://doi.org/10.1017/S0954579413000564>
- Rutter, M., & Uher, R. (2012). Classification issues and challenges in child and adolescent psychopathology. *International Review of Psychiatry*, 24, 514-529.
- Samuel, D., & Widiger, T. (2006). Clinicians ratings of clinical utility: A comparison of the DSM-IV and Five-Factor Models. *Journal of Abnormal Psychology*, 115, 298-308. doi: 10.1037/e594562007-001
- Samuel, D. B., & Widiger, T. A. (2011). Clinicians use of personality disorder models within a particular treatment setting: A longitudinal comparison of temporal consistency and clinical utility. *Personality and Mental Health*, 5, 12-28. doi: 10.1002/pmh.152.
- Sauer-Zavala, S. (in press). Measurement to improve treatment delivery: A commentary on the hitop measure development project. *Assessment*, <https://doi.org/10.1177/1073191121105095d>
- Scheiderer, E., Carlile, J. A., Aosved, A. C., & Barlow, A. (2017). Concurrent dialectical behavior therapy and prolonged exposure reduces symptoms and improves overall quality of life for a veteran with posttraumatic stress disorder and borderline personality disorder. *Clinical Case Studies*, 16(3), 216-233. <https://doi.org/10.1177/1534650116688557>

- Simms, L., Wright, A., Cicero, D., Kotov, R., Mullins-Sweatt, S., Sellbom, M., Watson, D., Widiger, T., & Zimmerman, J. (in press). Development of measures for the Hierarchical Taxonomy of Psychopathology: A collaborative scale development project. *Assessment*.
- Smith, B. P., Coe, E., & Meyer, E. C. (2020). Acceptance and commitment therapy delivered via telehealth for the treatment of co-occurring depression, PTSD, and nicotine use in a male veteran. *Clinical Case Studies*, 20(1), 75-91. <https://doi.org/10.1177/1534650120963183>
- Sprock, J. (2003). Dimensional versus categorical classification of prototypic and nonprototypic cases of personality disorder. *Journal of Clinical Psychology*, 59, 991-1014. <http://dx.doi.org/10.1002/jclp.10184>
- Stanton, K., Brown, M. F. D., Bucher, M. A., Balling, C., & Samuel, D. (2019). Self-ratings of personality pathology: Insights regarding their validity and treatment utility. *Current Treatment Options in Psychiatry*, 6, 299-311. <https://doi.org/10.1007/s40501-019-00188-6>
- Substance Abuse and Mental Health Services Administration (SAMHSA) (2020). Center for behavioral health statistics and quality. *Mental Health Annual Report: 2013-2018. Use of Mental Health Services: National Client-Level Data*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Tyrer, P. (2018). Dimensions fit the data, but can clinicians fit the dimensions? *World Psychiatry*, 17(3), 295-296. <https://doi.org/10.1002/wps.20559>
- Widiger, T. A., Sellbom, M., Chmielewski, M., Clark, L. A., Deyoung, C. G., Kotov, R., Krueger, R. F., Lynam, D. R., Miller, J. D., Mullins-Sweatt, S., Samuel, D. B., South, S. C., Tackett, J. L., Thomas, K. M., Watson, D., & Wright, A. G. C. (2018). Personality in a hierarchical model of psychopathology. *Clinical Psychological Science*, 7, 77-92. <https://doi.org/10.1177/2167702618797105>.
- Widiger, T. A., & Simonsen, E. (2005). Alternative dimensional models of personality disorder: Finding a common ground. *Journal of Personality Disorders*, 19, 110-130. <http://dx.doi.org/10.1521/pedi.19.2.110.62628>
- Zimmerman, M. (in press). Why hierarchical dimensional approaches to classification will fail to transform diagnosis in psychiatry. *World Psychiatry*, 20. <http://dx.doi.org/10.1002/wps.20815>

- Zimmerman, M., McGlinchey, J. B., Chelminski, I., & Young, D. (2007). Diagnostic co-morbidity in 2300 psychiatric out-patients presenting for treatment evaluated with a semi-structured diagnostic interview. *Psychological Medicine*, 38(2), 199-210.
<https://doi.org/10.1017/s0033291707001717>
- Zimmerman, M. (2016). A review of 20 years of research on overdiagnosis and underdiagnosis in the Rhode Island methods to improve diagnostic assessment and services (MIDAS) project. *The Canadian Journal of Psychiatry*, 61(2), 71-79.
<https://doi.org/10.1177/0706743715625935>
- Zimmerman, M., Balling, C. Dalrymple, K., & Chelminski, I. (2019). Have treatment studies of depression become even less generalizable?: Applying the inclusion and exclusion criteria in placebo controlled antidepressant efficacy trials published over 20 years to a clinical sample. *Psychotherapy and Psychosomatics*, 88, 165-170.
<https://doi.org/10.1159/000499917>

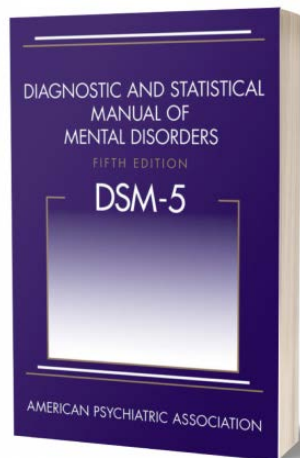
APPENDIX A

Introduction to Diagnostic Models in Study Survey

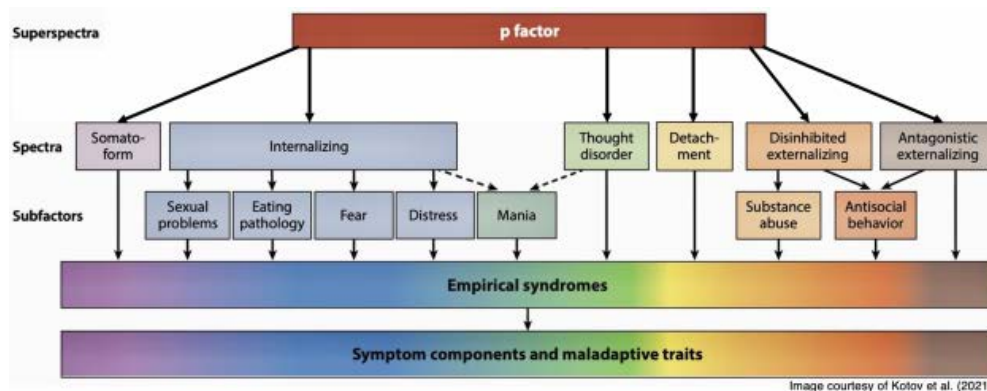
For this survey, you will read a brief clinical vignette about a client presenting for an outpatient intake. You will then diagnose the client according to two diagnostic systems.

One is the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), and the other is the Hierarchical Taxonomy of Psychopathology (HiTOP).

In the United States, the DSM-5 is *the* reference for the diagnosis and categorization of all mental disorders. It has been translated into 24 languages and has provided a common language for researchers, clinicians, educators, students, policy-makers, and third-party reimbursement entities to communicate with one another.



HiTOP was introduced in 2017 as an alternative to the DSM and a system for dimensionally classifying all variants of psychopathology. This model conceptualizes the structure of psychopathology as a hierarchy; at the bottom of the hierarchy are psychiatric signs and symptoms, and at the top are increasingly broad dimensions such as internalizing and externalizing factors.

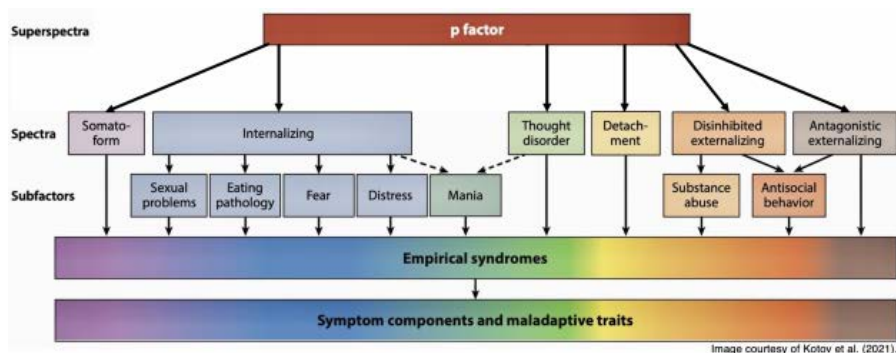


APPENDIX B

HiTOP Diagnostic Method

You will now rate the client according to the HiTOP approach.

- You will begin at the top of the hierarchy (i.e. the general psychopathology superspectrum) and work downward.
- You will only be asked to score the constructs that fall beneath the umbrella of difficulties that you identify as "somewhat" or "very much" a problem on the previous level of the hierarchy.



Superspectrum: General Psychopathology. Similar to the g factor of general intelligence, this construct represents the broad level of psychopathology severity and impairment, including all psychological disorders.

Spectra: Internalizing. This includes internalized symptoms of psychopathology such as anxiety, depression, fear, distress, sexual problems, eating problems, or certain mood components of mania.

Spectra: Thought Disorder. This includes symptoms such as disorganized thought and speech, reality distortion, cognitive and perceptual dysregulation, unusual beliefs and experiences, avolition, psychosis, or components of mania.

Spectra: Externalizing, Disinhibited. This includes problems such as substance use, aggression, excitement-seeking, impulsivity, low empathy, irresponsibility, or rebelliousness.

Spectra: Externalizing, Antagonistic. This includes trait-level difficulties such as attention-seeking, dominance, callousness, grandiosity, or manipulativeness.

Spectra: Detachment. This includes trait-level difficulties with depressivity, anhedonia, social withdrawal, or suspiciousness

Internalizing Subfactor: Distress. This subfactor of internalizing includes symptoms such as depressiveness, worry, affective trauma response, or general affinity for distress.

Internalizing Subfactor: Fear. This subfactor of internalizing includes symptoms such as obsessions, compulsions, phobias, or concern for panic attacks.

Internalizing Subfactor: Eating Problems. This subfactor of internalizing includes symptoms such as restricting, bingeing, purging, or preoccupation with eating, weight, or body size.

Internalizing subfactor: Sexual Problems. This subfactor of internalizing includes symptoms such as low sexual desire, difficulties with orgasm or arousal, or sexual pain.

Thought Disorder Subfactor: Mania. This subfactor of thought disorder includes symptoms such as euphoric activation, hyperactive cognition, or reckless over-confidence.

Externalizing, Disinhibited Subfactor: Substance use. This subfactor of externalizing, disinhibited includes problematic use or misuse of substances such as alcohol or illicit drugs.

Externalizing Subfactor: Antisocial Behavior. This is a subfactor of both the disinhibited AND antagonistic externalizing spectra. It includes symptoms such as aggression, boredom proneness, low honesty, psychopathy, or irresponsibility.

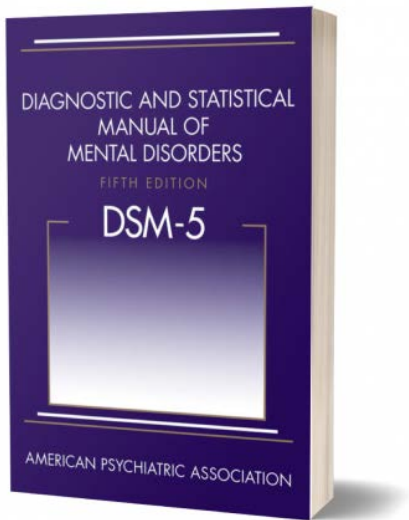
APPENDIX C

DSM Diagnostic Method

Please rate the client according to the following DSM-5 diagnoses.

- You are welcome to use a DSM for this step.
- You can make multiple selections in a category.
- Ignore the categories you deem irrelevant to the client's presentation.

The following lists are organized to reflect the table of contents of the DSM-5. Not all specifiers are included.



EXAMPLES OF LIST ORGANIZATION:

Schizophrenia Spectrum and Other Psychotic Disorders

- ☐ Schizotypal (Personality) Disorder
- ☐ Delusional Disorder
- ☐ Brief Psychotic Disorder
- ☐ Schizophreniform Disorder
- ☐ Schizophrenia
- ☐ Schizoaffective Disorder Bipolar type
- ☐ Schizoaffective Disorder Depressive type
- ☐ Substance/Medication-Induced Psychotic Disorder
- ☐ Psychotic Disorder Due to Another Medical Condition
- ☐ Catatonia Associated With Another Mental Disorder
- ☐ Catatonic Disorder Due to Another Medical Condition
- ☐ Unspecified Catatonia
- ☐ Other Specified Schizophrenia Spectrum and Other Psychotic Disorder
- ☐ Unspecified Schizophrenia Spectrum and Other Psychotic Disorder

Anxiety Disorders

- ☐ Separation Anxiety Disorder
- ☐ Selective Mutism
- ☐ Specific Phobia
- ☐ Social Anxiety Disorder (Social Phobia)
- ☐ Panic Disorder
- ☐ Agoraphobia
- ☐ Generalized Anxiety Disorder
- ☐ Substance/Medication-Induced Anxiety Disorder
- ☐ Anxiety Disorder Due to Another Medical Condition
- ☐ Other Specified Anxiety Disorder
- ☐ Unspecified Anxiety Disorder

APPENDIX D

Vignette A

Mia is a 32-year-old woman presenting for an outpatient intake appointment. She is married with two children and employed full-time. Mia is the primary care provider for her two daughters (ages 5 and 7), and her husband of nine years is deployed by the military. Mia served in the military herself from ages 18 to 25, and she has worked as an administrative assistant for a software company in the years since. She has no close relatives in the area and limited social support.

Mia presents for treatment with complaints of panic attacks, agitation, poor concentration, racing thoughts, isolative behavior, and poor sleep. Mia describes persistent ruminative worry throughout the day as well as experiencing panic attacks about once per week. She reports finding it markedly difficult to focus on tasks—even on straight-forward responsibilities such as paying bills. She identifies self-consciousness over feeling startled and distracted by most noises and other people. Despite persistent fatigue, Mia describes waking frequently throughout the night and not being able to go back to sleep, sometimes waking for the day as early as 3 or 4am. Mia describes herself as “worthless” and reports feeling great shame when sharing stories about herself and her past behaviors. She also identifies feelings of loneliness that conflict with her fear and avoidance of initiating social interaction with others, noting that the vast majority of her contact with other adults takes place through surface-level interactions at work. She reports ongoing daily use of alcohol and cannabis, and she confirmed that she turns to these substances as a means of managing difficult emotions.

Mia describes a significant trauma history including childhood sexual abuse (CSA) and physical abuse, intimate partner violence, military sexual trauma, and traumatic losses of loved ones. She reports sexual abuse by her step-grandfather from age 10 to 12, and physical abuse by her stepfather from age 13 to 17. Mia also reports that she and her husband have been physically aggressive toward one another on multiple occasions, though not to the point of requiring medical attention. Finally, Mia enlisted in the military at age 18 and reports experiencing unwanted sexual advances, sexual harassment, and inappropriate touching by a superior officer. Although she attempted to report this behavior, she did not feel her allegations were taken seriously until the individual was reported by an additional woman. Mia notes efforts to avoid distressing thoughts, feelings, and memories of her traumatic experiences, and prolonged, marked physiological reactions to reminders of her traumas. She also notes having difficulty with intimacy due to her history of sexual abuse.

Records indicate that Mia has a history of two serious suicide attempts in adolescence after which she reportedly received inpatient psychiatric care. Mia also describes a history of non-suicidal self-injury (NSSI) occurring in adolescence, remitting, and then resurfacing in the year prior to presenting for treatment. Reporting across her adult life, Mia endorses NSSI of cutting and burning her wrists and thighs, as well as other potentially risky behaviors such as driving under the influence, over-eating then forcing herself to vomit, and, when her husband is deployed, indiscriminate sexual encounters with strangers. The frequency of these behaviors appears to

coincide with the intensity of her life stressors. Upon presentation for treatment, Mia reported cutting her thighs with a razor approximately three times per month.

According to Mia, these various stressors and psychiatric symptoms have negatively impacted her performance as an administrative assistant. This has amounted to four disciplinary meetings with her supervisor over the past six months due to a lack of productivity and frequently calling out of work on short notice.

Vignette B

Mia is a 32-year-old woman presenting for an outpatient intake appointment. She recently separated from her spouse of five years and now lives with a roommate while working full-time in the customer service industry. She frequently interacts with her younger sister as well as her roommate, but she otherwise has no relatives in the area and limited social support. Mia denies being in a romantic relationship but reports engaging in unprotected sex with strangers she meets online about twice per week.

Mia presents for treatment due to ongoing difficulties with weight and body image, use of cocaine and alcohol, and “a lifetime of sadness.” Mia describes nightly episodes of eating large quantities of food during which she searches for food as if on autopilot and eats rapidly until she feels “stuffed.” Mia reports attempting to counteract her nighttime eating with intense cardio exercise for several hours at a time about four times per week. This standard often requires her to wake up at 4am to go to the gym before work, despite feeling exhausted. Mia reports vacillating between feeling depressed, worthless, irritable, and angry throughout the day most days. Her bad moods last from “minutes to days.” She also describes herself as a “worry wart” and requires daily support from her sister to assuage her various fears throughout the day. Mia adds that she has a very difficult time relaxing because the feelings of nervousness and dread are so distracting. This array of difficult emotions often manifests through somatic symptoms such as crying, shortness of breath, and nausea. Mia notes that sometimes, eating or drinking alcohol are the only relief from her emotional pain. Mia also reports financial difficulties that stem in part from “spending sprees” in which she spends so much money on food that she does not have enough to pay her other bills. According to Mia, these various stressors and psychiatric symptoms have negatively impacted her performance at work, leading to four disciplinary meetings with her supervisor over the past six months due to a lack of productivity and frequently calling out of work on short notice. A recent written evaluation from her boss indicated that organization, focus, and patience are areas of weakness for Mia.

Mia describes a complicated history with cocaine and alcohol use. Mia first used cocaine at age 16, after the drug was supplied to her by her mother. From then on, she used cocaine about twice per week—frequently with her mother—until about five years ago; now, she uses on occasion when she is intoxicated from binge drinking and needs a “pick-me-up” from a low or tired mood. Mia is aware that the cocaine use is interwoven with her eating habits because the cocaine suppresses her appetite. Regarding alcohol use, Mia reports consuming two to three standard alcoholic beverages about five times per week, as well as consuming more than 15 drinks in one day about once per week.

Mia believes many of her difficulties are related to her mother's criticisms of Mia's eating habits, weight, and body shape throughout her life. She describes experiencing poor body image and "a toxic relationship with food" since the age of 7; over the years, she has reportedly altered between limiting her caloric intake, or over-eating paired with over-exercise or laxative use. Mia recalls first feeling chronic low mood during her college years. While she does believe this onset in part because of her difficulties with eating and weight, she also identifies her perceived lack of relationships and accomplishments as causes.

Records indicate that as an adult, Mia has participated in several residential, intensive outpatient program (IOP), and partial hospitalization programs to treat her difficulties with eating, cocaine use, and alcohol use. She was most recently in a substance-use-focused IOP following a relapse with alcohol use. During these intensive programs, her substance use substantially reduces or stops altogether. Mia notes that she is prone to relapse upon stepping down to outpatient care.

Vignette C

Mia is a 32-year-old woman presenting for an outpatient intake appointment. She is single and lives alone. Following over a decade of service in the US army, she is enrolled in online undergraduate classes full-time and holds a part-time security job. Although she reports having positive relationships with friends and family, she has been socially isolated over the past six months. Specifically, she states that for most of her adult life, she communicated with close friends and family weekly, but she has had almost no contact with anyone outside of work in the past half-year. She also describes a disinterest in romantic relationships as of late and has not been on a date in many months.

Mia presents for treatment with complaints of difficulty focusing, decreased motivation, and chronic feelings of nervousness, as well as a desire to quit smoking cigarettes. Mia describes daily, persistent feelings of stress and tension throughout her body that make it "impossible" for her to relax. She reports feeling constantly distracted or startled by benign sound or movement in her environment. Mia also describes unusual procrastination that has significantly impaired her ability to keep up with her responsibilities over the past six months. She identifies sleep as a primary form of procrastination for her; she describes sleeping 2-3 hours more per night than is typical for her, as well as napping for 1-2 hours during the day. Mia reports being almost entirely inactive for two to three days per week due to feeling extremely tired and unmotivated. She spends these days in bed, sleeping, watching TV, or scrolling through social media, and she only leaves bed to retrieve snacks—most often junk food—from the kitchen. She adds that her ongoing difficulties with thinking and concentrating only seem to intensify on such days, and she finds herself constantly switching between television shows and social media platforms. On these particularly difficult days, Mia occasionally fantasizes about dying, such as about receiving a terminal cancer diagnosis or being involved in a fatal car accident, but she has never had plans or made an effort to hurt or kill herself. Mia thinks the strain of these various difficulties is compounded by guilt and shame for not living up to her expectations for herself—particularly because her expectations are rooted in the high standards for conduct and achievement she developed in the military.

Mia describes a highly stressful and difficult period of military training at age 20, including frequent hazing. Three years into her military career, she witnessed the suicide of a fellow service

member and close friend via firearm. Records indicate that Mia was ordered to receive psychotherapeutic intervention at the military base after the event. Mia states that the intervention was helpful overall, and because of it, the trauma gradually “stopped consuming [her] life.” Nonetheless, Mia continues to experience intrusive memories of what happened and efforts to avoid thoughts related to the trauma.

Mia began smoking cigarettes early in her military career, around age 21. What began as an effort to assimilate to her social environment escalated to smoking about half a pack of cigarettes per day from ages 23 to 30. In the two years since leaving the military, Mia transitioned to vaping and has used decreasing levels of nicotine in an effort to quit. She reached the lowest possible dose of nicotine about five months ago, but she has not been able to quit completely. She experiences urges to use her vape throughout the day nearly every day, and she finds she is unable to focus on tasks unless the urges are satisfied.

Mia struggles to identify the onset of these mental health problems, but she believes they only began to interfere with her life following the suicide of her colleague. She notes the recent 10-year anniversary of this death could be related to her more recent struggles, but she thinks her primary stressor is “not knowing what [she] wants to do with [her] life.” After two years enrolled as an undergraduate, she does not feel she has made progress in finding a major that interests her. According to Mia, these various stressors and psychiatric symptoms have negatively impacted her ability to complete her work and daily tasks. This has led to disciplinary meetings with her academic advisor over the past six months due to falling grades and frequently missing class.

APPENDIX E

Optional Free-Response: What further thoughts do you have on these topics or this survey?

COMMENTS ABOUT THE DSM

DSM V is awful.

Honestly as I've gone on in practice the dsm offers less and less utility and seems only relevant as a requirement.

I appreciate introduction to an alternative diagnostic system. I am frequently frustrated by the limits of the DSM in my work and find that multi-disciplinary teams on which I work and in communication with other providers a great deal of time is spent discussing/debating diagnosis even though we often all note the same presenting problems/symptoms.

The DMS 5 is ridiculous. The diagnoses are so specific (too much minutia) that it's hard to apply to real patients. It's not practical or useful clinically. The current diagnostic criteria used for PTSD is not credible at all. every person on earth now meets the criteria for PTSD (minus the significant impairment which unfortunately many clinicians ignore). In addition, they added physical aggression to the diagnostic criteria. Anybody who has treated a significant number of patients with PTSD will tell you that this is NOT true. I've treated thousands and only a few have been violent. Those patients had personality disorders or they were intoxicated. Now with this being part of the diagnostic criteria, violent people have an excuse for their behavior. The only patients who are violent because of true mental illness is psychotic patients, and even that is very rare. Otherwise it's straight up cluster B behavior or substance abuse. The DSM 5 is designed / appropriate for research rather than clinical use. The majority of the world uses ICD rather than DSM. In addition, there's huge money made from the DSM system. I'm clearly not a fan & many of my colleagues feel the same way. I'm interested in what your results will be from this survey =)

The DSM is required, my choice in the matter is fairly irrelevant unless the system changes. I follow many levels of requirements as part of a large system seeing Medicare and Medicaid and DSM diagnosis is a integral part of being able to get clients the services they need. I would also like to mention the case study technically appeared to meet criteria for many DSM diagnosis, including MDD, but I chose PTSD specifically because it includes mood, social and somatic components and better reflected the treatment they might need. I see some utility in the other system and would not discourage further exploration of it's use.

I would very much like to see an alternative to DSM for diagnosing. The DSM artificially divides symptoms into separate diagnoses that do not really capture the individual's experience and problems. It also encourages the development of empirically supported treatment protocols that target only ONE diagnosis or symptom, while ignoring the whole person's existential, physiological, and social concerns.

COMMENTS ABOUT HITOP

I didn't like that I was forced to choose between the DSM and Hi-Top as which diagnostic classification method I would use. I would have liked an option of "I'm not sure." I am curious about how HiTop results are reported. I think if I would have known how the results were reported, I may have had a different response regarding which system I prefer. I think the HiTop is lacking in some face validity. I appreciate its attempt to be broader and I think it sacrifices some specificity in communication because of that.

I like the idea of this newer approach, but I didn't fully understand it, especially the rating part at the bottom. Also, it seemed that some constructs were missing or in the wrong place. One fell under both umbrellas and I wondered why.

I may need to learn more about it

I prefer dimensional systems and more experience near systems for diagnosis

I think it would be hard to use the HITOP system to communicate to colleagues because I don't know anyone else in practice who uses it or has said they know about it.

I think the HiTop and similar models are the future of diagnosis and treatment. There is so much overlap between the DSM-5 Syndromes and high rates of co-morbidity--the HiTop is consistent with the science and clinical practice. It will, however, be a difficult shift for providers and we'll have to come up with some creative ways to translate HiTop assessments for our patients/clients. I don't think it will be too difficult since we already have assessments that rate people on different spectrum like the Achenbach and MMPI systems.

I was totally unfamiliar with the hierarchical system, but found it intuitive enough to apply to the vignette on first pass.

Interesting! I'd like to learn more about HiTop.

So glad someone is doing research in this important area! Would love to see the HiTOP model being used and really liked the way in which it was presented here.

My preference for the DSM is largely based on the fact that i work in an anxiety specialty setting and I find that knowing the specific subtypes of anxiety and OCD as laid out in the DSM has important treatment implications for designing tailored exposure-based CBT exercises to support the client most effectively.

I work with a lot of neurodiverse individuals and I don't see some of that clearly represented on the HiTOP in ways that would be meaningful for my assessment clients. From a general individual therapy perspective, I like the HiTOP quite a bit, as it seems functionally focused. I don't love that ADHD is thrown in with antisocial, as that seems to focus on problematic impulsive behaviors but misses the broader difficulties of executive functioning that still lead to very real functional impairments that I don't think are wonderfully captured here.

COMMENTS COMPARING THE DIAGNOSTIC SYSTEMS

HiTOP, while more conceptually interesting than the DSM-5, does not more fully contextualize the client's presentation as does a functional analysis.

Although I believe the HiTop system to be a large improvement over the DSM-5; I find that functional contextual approaches are much more beneficial than either classification system for treatment planning. I am much more interested in the development of process-oriented diagnostic systems rather than categorically-based systems, such as the DSM and HiTop.

I love the idea of a dimensional approach to dx, but I don't know how to use HiTop. I think having a specific diagnosis is often helpful for patients, as it helps them understand that they are not alone and that there is a name for what they are experiencing. I think a dimensional approach is helpful for communicating between professionals, establishing a case conceptualization, and enhancing psychoeducation given to patients about their diagnoses.

The HiToP approach provided a better framework for sorting client symptoms/problems as compared to DSM, where I found myself spending more time thinking about the category that was most applicable (e.g., Is the worry + physical symptoms indicative of GAD, or is this more of an unspecified anxiety disorder; does it even matter given the likely personality pathology that drives things?). I do think something missing in both approaches is the limited attention to function (why do these symptoms arise? in what ways do they work for the client?).

My first impression is that your formulation seems more client-friendly, in terms of explaining problems to clients. I like that. But the DSM seems more clinician-friendly. This is also a function of my level of familiarity and years of DSM use.

The psychodynamic diagnostic manual has some similar methods for categorizing functioning in various domains as opposed to categorizing symptoms, which seems similar to the HiTOP model presented here. I like this model, though as I'm unfamiliar with it will likely continue to defer to the DSM and PDM for diagnosis until I can learn more. Thank you!

There are so many systems built around the DSM (/ICD), and while it's promising that NIMH is distancing from it, until there is a way to shift billing/insurance away from it, these remain academic questions.

The vignette was written to be someone who meets some criteria for a lot of different psychiatric disorders. As such, it seemed there was a bit of a demand characteristic to choose/favor HiTOP in answering questions. It would be more meaningful and generalizable to compare clinician favorability of diagnostic systems using multiple vignettes with descriptions of individuals across a range of complexity. Further, an opportunity was missed in this study to examine the treatment utility of each diagnostic system (One question was asked, but more could be in future studies). To wit, it could be helpful to assess clinician preference for type of intervention based on diagnostic system. It may be that HiTop is equally unhelpful as DSM-5 in differentially informing specific treatment plans. DSM-5 may be more useful for most clinicians in communication, imperfect as it is (and it is!). HiTop is a niche system that academics may be aware, but it routine real-life clinics across the US, this system has not made any real inroads. It's a huge challenge, and I appreciate you doing this survey as part of efforts to determine how to address it.

This is really interesting as I have never heard of this approach per se, but is consistent with the psychiatric (psychosocial) rehabilitation approach used prior to VA when working with folks with severe SMI in long-term state hospital program - the same problem behaviors can be seen in individuals with very diverse DSM diagnoses but it's the problem behaviors not the dx that keep a person committed and hospitalized rather than successfully living in the community, even in group home.

this is great. i appreciate these alternate diagnostic models and find them more useful than the DSM, whose only value for me is getting the proper code for insurance and flattening out wrinkly papers.

Working in primary care I need my notes to be interpretable by the referring provider and the medical team. This questionnaire did not assess the ease with which my notes could communicate with others on the patient's treatment team. In my setting, Hitop would be confusing for the providers (but they are very aware of DSM diagnoses and will make treatment decisions based on that). Also, eligibility to many social services or disability claims hinge on either a DSM or ICD diagnosis and using Hitop would likely be unhelpful for the patient to access these resources. I like hitop better, sometimes I use it for case conceptualization, but within our medical system it's hard to see how it could be meaningfully implemented. I would love it if we could! Best of luck with your research

Would love to have a better diagnostic tool so great work on this!

COMMENTS ABOUT THE STUDY

For the question about which method of diagnosing/conceptualizing one prefers, I believe it would be helpful to have an option along the lines of "unsure" or "need more information/training."

Having participated in Hi-TOP research, I felt biased and tried to hold that bias back while completing the survey.

I found the hiTOP algorithm confusing.

It would have been nice to have had more training on HiTOP before being asked to use it.

I'd be curious to see the results of the study!

It wasn't clear whether we were going by potential diagnoses or actual diagnoses with the DSM so I erred towards all that might apply.

It would have been helpful to go back and re-read the vignette when doing the diagnoses. In normal circumstances you can review your notes if needed when formulating. I had to formulate my diagnoses based on memory and any inaccuracies may have been rectified if a re-review of the vignette had been possible.

Very interesting and important research! I look forward to the results of your study!

My window wouldn't open so had to do the vignette from memory, sorry! I had some hesitation with DSM diagnoses as I couldn't remember exact details (e.g., frequency of binge/purge episodes for diagnosis, content/triggers for panic attacks)

important but tricky work! I'm mostly work with kids so the presenting problems described here are somewhat beyond what I typically see. I'm less familiar with eating disorders and substance use disorders, and I mostly skimmed the vignette and made my choices quickly. My biggest takeaway is that these patient was elevated on a number of domains, which would make both diagnosis and treatment planning challenging to figure out what to prioritize.
