

Reconstructing Psychopathology: A data-driven reorganization of the symptoms in DSM-5

Miriam K. Forbes, School of Psychological Sciences, Macquarie University
ORCID: 0000-0002-6954-3818

Andrew Baillie, Sydney School of Health Sciences, Faculty of Medicine & Health, University of Sydney & Sydney Local Health District, ORCID: 0000-0003-3424-5703

Philip J. Batterham, Centre for Mental Health Research, College of Health and Medicine, The Australian National University, ORCID: 0000-0002-4547-6876

Alison Caelear, Centre for Mental Health Research, College of Health and Medicine, The Australian National University, ORCID: 0000-0002-7028-725X

Roman Kotov, Department of Psychiatry, Stony Brook University, ORCID: 0000-0001-9569-8381

Robert F. Krueger, Department of Psychology, University of Minnesota, ORCID: 0000-0001-9127-5509

Kristian E. Markon, Department of Psychological and Brain Sciences, University of Iowa, ORCID:

Louise Mewton, Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, ORCID: 0000-0002-7812-296X

Elizabeth Pellicano, Psychology and Language Sciences, University College London, ORCID: 0000-0002-7246-8003

Matthew Roberts, School of Computing, Macquarie University, ORCID: 0000-0002-2553-6157

Craig Rodriguez-Seijas, Department of Psychology, University of Michigan, ORCID: 0000-0003-4595-6658

Matthew Sunderland, Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, ORCID: 0000-0001-8452-364X

David Watson, Department of Psychology, University of Notre Dame, ORCID: 0000-0002-9605-0576

Ashley L. Watts, Psychological Sciences, Vanderbilt University, ORCID: 0000-0002-8210-4748

Aidan G.C. Wright, Department of Psychology and Eisenberg Family Depression Center, University of Michigan; ORCID: 0000-0002-2369-0601

Lee Anna Clark, Department of Psychology, University of Notre Dame
ORCID: 0000-0002-2571-4347

Correspondence to: Miriam K. Forbes, PhD
Australian Hearing Hub, Room 3.803
School of Psychological Sciences,
Macquarie University,
Sydney Australia 2109
Email: miri.forbes@mq.edu.au

This pre-print has not been peer reviewed

Author Contributions: Authors Baillie to Wright are listed alphabetically. Forbes was responsible for conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, visualization, and writing – original draft. Roberts led the development of the survey plug-in that facilitated recruitment and retention (software and methodology). Baillie, Batterham, Calear, Clark, Forbes, Krueger, Markon, Mewton, Pellicano, Rodriguez-Seijas, Sunderland, Watson, and Watts worked on developing the items assessing DSM-5 symptoms (methodology). Baillie, Clark, Forbes, Kotov, Krueger, Markon, Sunderland, Watson, and Wright worked on the a priori analytic plan (methodology). Baillie, Batterham, Calear, Clark, Kotov, Markon, Mewton, Pellicano, Rodriguez-Seijas, Sunderland, Watson, Watts, and Wright worked on interpreting the results. All authors contributed to review and editing of the manuscript.

Conflicts of Interest: The authors declare that there were no conflicts of interest with respect to the authorship or the publication of this article.

Acknowledgements: The authors would like to thank Tim Slade and Lorna Peters for their input on the development of the self-report item battery; Yael Perry and Ashleigh Lin for feedback on the eight gender dysphoria items; Jaclyn Shepard and Lee Ritterband for feedback on the six elimination disorder items; Cele Richardson for feedback on the sleep-wake disorder items; Owen Forbes for feedback on the suicide-related items; Brier Michelsen and Hannah Woodbridge for their work on splitting *DSM* criteria into their constituent symptoms; Anton Harris, Katherine Faure, and Bryan Neo for their work on qualitative content coding of *DSM* symptom repetition; Brier Michelsen, Katherine Faure, and Maddison Twose for their project management work as research assistants during data collection; Emma Walker, Tam Du, Oscar Gardiner, and Kyle Long for their work getting the survey visualization plug-in started; the community organizations who helped us to advertise the study; and the study participants for their time completing the survey!

Funding: This study was funded by Investigator Grant from the Australian National Health and Medical Research Council (grant number GNT1194292).

Abstract

In this study, we reduced the *DSM-5* to its constituent symptoms and reorganized them based on patterns of covariation in individuals' ($n = 14,762$) self-reported experiences of the symptoms to form an empirically derived hierarchical framework of clinical phenomena. Specifically, we used the points of agreement among hierarchical principal components analyses and hierarchical clustering, as well as between the randomly split primary ($n = 11,762$) and hold-out ($n = 3,000$) samples, to identify the robust constructs that emerged to form a hierarchy ranging from symptoms and syndromes up to very broad superspectra of psychopathology. The resulting model had noteworthy convergence with the upper levels of the Hierarchical Taxonomy of Psychopathology (HiTOP) framework and substantially expands on HiTOP's current coverage of dissociative, elimination, sleep-wake, trauma-related, neurodevelopmental, and neurocognitive disorder symptoms. We also mapped some exemplar *DSM-5* disorders onto our hierarchy; some formed coherent syndromes, whereas others were notably heterogeneous.

Reconstructing psychopathology: A data-driven reorganization of symptoms in DSM-5

By delineating the constructs we study and treat, classification of psychopathology drives both clinical research and practice. For decades, the diagnoses defined in the traditional classification systems like the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* and *International Classification of Diseases (ICD)* have shaped the field, and have facilitated much progress. However, a growing body of research highlights limitations with traditional mental disorder diagnoses that persist in the *DSM-5-TR* (American Psychiatric Association [APA], 2022). Some prominent limitations include: 1) heterogeneity within diagnoses due to its ‘checklist’ approach, which allows two people with the same diagnosis to have no symptoms in common (e.g., Fried & Nesse, 2015); 2) fuzzy boundaries between diagnoses due to symptom overlap that inflates their surface similarity (e.g., Forbes et al., 2023); 3) misalignment with clinical symptom profiles (e.g., Newson et al., 2021); and correspondingly, 4) low inter-rater reliability of many diagnoses (e.g., Regier et al., 2013). Many now believe that we have reached the point where *DSM*-defined constructs are hindering rather than facilitating progress in understanding mechanisms and improving treatment outcomes, and new approaches to conceptualizing psychopathology are therefore gaining momentum (e.g., Cuthbert & Insel, 2013; Hofmann & Hayes, 2019; Lilienfeld & Treadway, 2016; McGorry et al., 2006).

This study aims to overcome the limitations of traditional diagnostic constructs and to bridge the growing divide between the massive research portfolio focused on *DSM* constructs and the new approaches to conceptualizing psychopathology that are moving away from these constructs. Specifically, this study aims to reorganize the symptoms described in the *DSM* into a new model of empirically derived constructs based on quantitative patterns in individuals’ experiences. By starting with the same pool of symptoms as the *DSM*—which captures decades of rich clinical observation on how psychopathology is manifested—this approach offers the advantage of facilitating a direct translation of the body of knowledge framed by traditional diagnoses.

Quantitative Approaches to Classification

Quantitative approaches to classification have gained prominence as the field has sought alternative approaches to traditional diagnostic categories and systems, culminating in the Hierarchical Taxonomy of Psychopathology (HiTOP), which aims to overcome the limitations of the traditional diagnostic categories and systems (Kotov et al., 2017, 2021). The HiTOP framework (see Figure S1) is a hierarchy of empirically derived dimensions with very broad ‘superspectra’ at the top—such as externalizing psychopathology—and increasingly narrow dimensions nested on each level moving down the hierarchy. These dimensions are quantitatively derived from patterns of covariation or comorbidity among symptoms and disorders and appear promising as an alternative approach to classification (see Conway et al., 2019; Kotov et al., 2021, for detailed reviews). However, limitations of the HiTOP framework restrict its ability to advance the field.

Limitations of HiTOP

Because much of the work underpinning HiTOP was based on analyses of the covariation or comorbidity among *DSM*-defined disorders (Kotov et al., 2017), some of the limitations of *DSM* diagnoses are baked into the HiTOP structure, which is ironic for a framework aiming to overcome those limitations. As one example, symptom overlap between diagnoses prevents us from disentangling which symptoms are truly nonspecific and transdiagnostic versus those that “belong” with a specific symptom set when allowed to covary freely (e.g., Stanton, 2020; Stanton et al., 2023). For instance, in an exploratory symptom-level analysis of a variety of self-report measures, Forbes et al. (2021) found that items assessing difficulty concentrating from measures of generalized anxiety disorder (GAD), major depressive disorder (MDD), and attention-deficit/hyperactivity disorder (ADHD) cross-loaded on all three of these constructs, whereas items assessing irritability from measures of anger, GAD, and mania loaded only with anger. Allowing symptoms to freely form empirically derived homogeneous constructs could also address heterogeneity within diagnoses, fuzzy boundaries between diagnoses, improve reliability, and answer recent calls for more precision in psychiatric phenotypes to advance neuroscience and psychiatric genetic research (e.g., Derks et al., 2022; Tiego et al., 2023; Watts et al., 2023b).

A related limitation of the HiTOP framework is that the detailed levels of the model are not well fleshed out. The official HiTOP model (Figure S1) currently includes a level of homogeneous symptom components and maladaptive traits, all of which represent subscales of existing self-report measures (Kotov et al., 2017). In one sense, most of these subscales were empirically derived in the process of each measure’s development. However, the measure development process typically involves dropping items that do not strongly define a single construct (e.g., that cross-load on multiple constructs or have weak loadings on the constructs of interest; see Clark & Watson, 1995, 2019), which eliminates interstitial and transdiagnostic symptoms and limits our understanding of an empirically derived structure of psychopathology. This characteristic of many symptom measures in psychology has also limited nearly all symptom-level studies of the structure of psychopathology, together with the common methodological feature of presenting items in predefined blocks that anchor responses to the target construct and introduce order effects (see Forbes et al., 2021 for a review). Further, the homogeneous symptom components and maladaptive traits in the HiTOP model vary substantially in breadth (e.g., alcohol problems are multidimensional; Watts et al., 2021), and some are closely related to *DSM* diagnoses—for example, the Interview for Mood and Anxiety Symptoms included the *DSM*-defined symptoms sets for MDD, GAD, and PTSD as part of the initial item pool in the measure development, and all of the resulting subscales are included together under the distress subfactor in HiTOP (Kotov et al., 2015, 2017; Watson et al., 2012).

Using the scope of the *DSM-5* (APA, 2013) as a point of reference, the breadth of coverage of psychopathology in HiTOP is also limited (e.g., mapping 8 of the 19 chapters in full, and 6 in part; Kotov et al., 2021). Whole classes of disorders are not yet part of the framework (e.g.,

neurocognitive, dissociative, and paraphilic disorders) and only subsets of disorders are included from certain chapters (e.g., neurodevelopmental, feeding and eating, obsessive-compulsive and related, trauma- and stressor-related, and disruptive, impulse-control and conduct disorders). If HiTOP is intended as an alternative classification system to the *DSM*, then it should aim to provide the same breadth of construct coverage (First et al., in prep).

Further, the HiTOP model itself is based on a schematic representation of the literature reviewed in Kotov et al. (2017), rather than an overarching statistical model that has been derived from or fit to data. The upper levels of the model were recovered in a meta-analysis spanning five of the six core spectra in the model (Ringwald et al., 2023), but this evidence, too, is based on *DSM*-defined diagnoses. Although there is a substantial body of evidence for the structural and external validity of the broad dimensions in the HiTOP framework (e.g., Kotov et al., 2020; Krueger et al., 2021; Watson et al., 2022), even these dimensions may be artificially shaped by features such as symptom overlap within *DSM* diagnoses (Forbes, 2023a) as well as the specific indicators included in each of the studies reviewed to build the model (e.g., confining analyses to a single spectrum, the emergence of bloated specifics due to disproportionate coverage of a single domain, or precluding dimensions from emerging due to insufficient coverage of relevant domains; Watts et al., 2021). To advance our understanding of psychopathology and improve treatment approaches, the field needs to progress toward an alternative classification system that is supported by more direct and specific indicators of psychopathology and not merely a reorganization of *DSM*-defined diagnoses (Levin-Aspenson, 2023). Overall, although HiTOP holds promise, its limitations are currently preventing it from fulfilling that promise.

The Present Study

The present study aims to address limitations of both the *DSM* and HiTOP by reorganizing the existing elements of the *DSM-5* into a hierarchy of empirically derived phenomena based on their patterns of covariation in individuals' self-reported experiences of the symptoms. By starting with symptoms as the units of analysis, we can 1) untether the outcome from the limitations of traditional diagnoses, 2) cover the full breadth of psychopathology described in the *DSM-5*, and 3) flesh out a comprehensive set of narrow homogeneous constructs that may pave the way to untangling links between risk factors, mechanisms, and treatment outcomes with psychopathology.

Transparency and Openness

The analysis plan for this study was preregistered on the Open Science Framework (OSF; <https://osf.io/uek39/>). There was one deviation from the analytic plan in the preliminary analyses, as noted in the Methods. Given the complexity and extent of the results, we have made the code and all output available on the OSF (<https://osf.io/gdvnv8/>) to facilitate alternative interpretations. The correlation matrices are posted privately on OSF (<https://osf.io/mc7z4/>) and will be shared publicly 12 months after the publication date of this paper. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All study procedures were approved by the Macquarie University Human Research Ethics Committee and

the project was conducted in accordance with the World Medical Association Declaration of Helsinki.

Method

Participants and Procedure

The online survey was conducted between December 2020 and October 2022, and was fully anonymous and confidential. Participants were recruited through social media advertising as well as with the help of a variety of Australian community groups, including groups with a focus on lived experience of mental illness (e.g., Blue Voices; GROW Australia), experiences of specific mental disorders (e.g., The Butterfly Foundation, which focuses on eating disorders and body image issues), and multicultural community groups (e.g., WellMob, focused on Indigenous Australians' mental health; Sydney Multicultural Community Services; Transcultural Mental Health Centre). Study advertisements used tag lines such as "Help us learn more about mental illness," "Help us learn more about mental health," and "Help create a detailed map of mental health[/illness]." The social media advertising strategy included targeted campaigns for men and for older adults (e.g., aged 60+) to reach demographic groups that are traditionally difficult to recruit into online surveys. The study was also made available to undergraduate students at Macquarie University to complete for course credit. We did not have an upper limit for a target sample size, but aimed to recruit a large and varied sample to maximize representation of low prevalence and severe symptoms; the sample size was ultimately determined by the exhaustion of resources for recruitment (i.e., the timeline for the project, and running out of funding for advertising costs as well as community groups to contact).

Of the 22,292 link clicks to participate in the survey, 11.4% did not begin the survey (specifically, 11.1% did not progress beyond the information and consent form; 0.3% did not consent to participate in the study), 1.2% were ineligible to participate due to being under 16 years of age—see supplementary materials for information on this inclusion criterion—and an additional 0.4% were detected as bots by Qualtrics based on V2 Completely Automatic Public Turing Tests [CAPTCHA] verification. Further, 19.5% dropped out in the first few pages of the survey before reaching questions about symptoms of psychopathology, and 1.2% were excluded for failing attention checks throughout the survey¹. Participants' IP addresses were not recorded, but Qualtrics prevented multiple survey submissions from the same IP address. Excluding these data left an analytic sample of 14,762 participants.

Table 1 describes the sociodemographic characteristics of the analytic sample in detail, which indicate substantial representation of people receiving treatment for their mental health, as well as substantial diversity of education levels, employment status, relationship status, residential-area type, and sexual orientation and gender identities. In contrast, a large majority (73.2%) of

¹People who saw 1 attention check had to pass it; people who saw 2-3 could fail 1; and people who saw 4+ needed to pass 75% of the attention checks they saw.

Table 1. *Sociodemographic information about the analytic sample (n = 14,762)*

Variable	N	%
Age in years (mean [SD] = 38.0 [18.01])		
Youth (16–24)	5,036	34.1%
Early adulthood (25–39)	3,056	20.7%
Middle adulthood (40–59)	4,274	29.0%
Later adulthood (60–91)	2,396	16.2%
Gender (all that apply)		
Women	8,119	55.0%
Men	5,279	35.8%
Non-binary/gender fluid	1,366	9.4%
Another identity not listed	170	1.2%
Not reported	1	0.0%
Transgender	720	4.9%
Highest level of education completed		
Less than primary/elementary school	46	0.3%
Completed primary/elementary school	2,072	14.0%
Completed high school/received GED	2,956	20.0%
Some college/university education	2,756	18.7%
Associate Degree or Technical Certificate	1,628	11.0%
Bachelor's Degree	2,385	16.2%
Some postgraduate education (e.g., started Master's or Doctorate Degree)	1,081	7.3%
Completed graduate degree (Master's or Doctorate)	1,835	12.4%
Not reported	3	0.0%
Employment status (all that apply)		
Employed full-time	3,741	25.3%
Employed part-time	3,313	22.4%
Unemployed	1,877	12.7%
Part-time student	752	5.1%
Full-time student	3,540	24.0%
Retired	1,474	10.0%
Homemaker or stay-at-home parent	410	2.8%
Self-employed	785	5.3%
Unable to work	1,447	9.8%
Another option not listed	661	4.5%
Racial identity or origin (all that apply)^a		
Asian	736	5.0%
Black or African American	128	0.9%
Hispanic or Latino	335	2.3%
Indigenous or Aboriginal	261	1.8%
Indian	128	1.2%
North African or Middle Eastern	185	1.3%
Pacific Islander	75	0.5%
White or Caucasian	10,813	73.2%
Another race or origin not listed	368	2.5%
Not reported	2,606	17.7%
Sexual orientation (all that apply)^a		
Straight or heterosexual	7,801	52.8%
Gay or lesbian	1,164	7.9%
Bisexual	2,447	16.6%
Another orientation not listed	1,139	7.7%
Not reported	2,608	17.7%
Relationship status (all that apply)^a		
In a relationship; living together or married	4,486	30.4%
In a relationship; not living together	1,701	11.5%
Widowed	239	1.6%
Separated	416	2.8%
Divorced	754	5.1%
Not in a relationship	4,765	32.3%
Another status not listed	232	1.6%
Not reported	2,593	17.6%
Country of residence^a		
Australia	8,438	57.2%
USA	3,562	24.1%
Other	171	1.2%
Not reported	2,591	17.6%
Residential area		
Urban area	3,549	24.0%
Suburban area	8,083	54.8%
Rural area	3,125	21.2%
Not reported	5	0.0%
Currently receiving treatment for mental health	9,410	63.7%
Ever received treatment for mental health	12,250	83.0%

Note.

^aThe variables with larger proportions of missingness (approx. 17%) were presented at the end of the survey as optional additional sociodemographic questions. GED: General Educational Development tests.

participants reported their racial identity as including being white or Caucasian—93% of whom ($n = 10,086$, 68.3% of the whole sample) endorsed white or Caucasian as their only racial identity—most (55.0%) participants identified as women, and most (57.2%) lived in Australia.

To maximize inclusivity, participants could opt to complete the mini (22.6% of participants selected this option), short (19.5%), medium (11.3%), or long (46.6%) versions of the survey with median (interquartile range) completion times of 13.3 (10.4–17.6) minutes, 23.3 (18.4–31.5) minutes, 39.9 (29.5–57.2) minutes, and 72.3 (47.8–124.6) minutes, respectively. The core survey items (described below) were fully randomised into 12 blocks: The mini survey included one of the 12 blocks, the short survey included three blocks, and the medium survey included six blocks; the blocks shown to each participant completing the mini, short, or medium length surveys were selected at random from all possible blocks. The long survey included all 12 blocks. Blocks were presented in random order and items within each block were also presented in a randomised order for all versions. Given the fully randomised item presentation and the planned missingness design, all participants who responded to even a single item were included in the final analytic sample and missing data were handled under a missing-at-random assumption.

The survey had AUD\$10,000² available of prepaid cash cards or charity donations allocated based on random draws: Participants could choose to opt into a cash draw or to donate their draw to a mental health charity (\$20, \$50, \$100, and \$250², respectively, for the varying survey lengths) in a separate survey that was not linked to their survey responses. Support resources for participants experiencing distress were provided throughout the survey. At the end of the survey, a personalized and interactive visualization showed each participant's data relative to the overall sample in terms of symptom severity, frequency, and patterns of covariation in the full dataset (see <https://helpuslearnmoreaboutmentalillness.webflow.io/visualisation-information>).

Measurement of *DSM-5* symptoms

Our analyses focused on the item pool developed to assess individuals' self-reported experiences over the past 12 months of all symptoms represented in the *DSM-5*.³ Diagnostic criteria for diagnoses in Section II of the *DSM-5* were reduced to a list of their constituent symptoms (i.e., subjective experiences of thoughts, feelings, behaviors, and physical symptoms), and criteria for Personality Disorders (Chapter 18) were replaced from the outset with a 100-item measure of the *DSM-5* Section III Alternative Model of Personality Disorders (i.e., the Personality Inventory for *DSM-5* [PID-5]; Krueger et al., 2012; Maples et al., 2015). The PID-5 short-form has been found to preserve the reliability and validity of the full PID-5 (Maples et al., 2015), which accounts for the reliable variance in Section II personality disorders (Anderson et al., 2014;

²Approximately \$6,565, \$13, \$33, \$66, and \$164 U.S. dollars, respectively.

³ Other measures included in the study but not used in the present analyses were the Kessler Psychological Distress Scale and the World Health Organization Disability Assessment Schedule. These will be included in the next planned project for these data.

Krueger & Markon, 2014) and maps onto the empirical structure of personality disorder criteria (Williams et al., 2018).

Signs and symptoms were excluded from further consideration if they required medical or specialized testing (e.g., IQ testing or polysomnography), could only be observed by others (e.g., cyanosis during sleep), were specific to a single culture or country (e.g., khyâl cap), or were only relevant to children (e.g., “is often truant from school”). We also separated symptoms from their causes and consequences, including associated distress and impairment (see Üstün & Kennedy, 2009), and removed descriptive information about symptom onset, duration, frequency, and severity. This list of 1516 symptoms was then coded for redundancy based on semantic and conceptual similarity, and redundant symptoms were collapsed into a single item (see Forbes et al., 2023b for more detail on the qualitative content coding method for redundancy).

The resulting 711 symptoms were then written in first person and past tense, remaining as close as possible to the phrasing of the *DSM-5* to ensure the original symptom was captured. This preliminary item pool underwent three rounds of readability testing and improvement, including readability calculators and two rounds of pilot testing with first-year undergraduate psychology students for feedback on which items were hard to understand, and why. This process resulted in rephrasing items to use simpler versions of technical words, restructuring sentences, and adding examples to symptoms, the latter of which were drawn from the *DSM-5* text in most cases. The revised items were sent to 16 experts in psychopathology measurement as well as content experts in specific domains of psychopathology (see Author Note) for 1) feedback on how well each item captured the corresponding symptom; 2) comments on any significant problems regarding the clarity, assumptions, knowledge/memory, and sensitivity/bias of the item; and 3) any suggestions for improving the wording of the item (Questionnaire Appraisal System; Wills & Lessler, 1999). Items were revised based on expert feedback and piloted a final time before launching the official survey. The final list of 680 items is listed in Table S1.

Participants reported how true or common each statement was for them in the past 12 months on a five-point scale from *Not at all true (Never)* to *Perfectly true (Always)*, after being instructed to think about their experiences across a wide variety of contexts. Participants were asked to select *Not at all true (Never)* if an item did not relate to their experience and no skip-structure was applied to the items. Details on the preprocessing of the data and first stage of dimension reduction are provided in the supplementary materials. Briefly, items with very low endorsement (<5% non-zero responses; $n = 24$ items) were pooled with other similar low endorsement items when possible ($n = 19$ items), or dropped if this was not possible ($n = 5$ items; see Table S1). Very highly correlated items ($\rho \geq .8$; $n = 16$ items) were combined by taking the mean of their standardised values. Further, items that were correlated $\rho < .3$ with every other item being analysed were excluded ($n = 13$ items; see Table S1).

Finally, the remaining 614 observed variables—covering 647 symptoms—were then subjected to both *iclust* (Revelle, 1979) and Ward's (1963) hierarchical agglomerative clustering.⁴ The convergence of the two clustering solutions was used to identify homogeneous composites to carry forward for analysis (i.e., only combining items into a cluster when both methods agreed). Using two different methods ensures there are no large clusters idiosyncratic to a specific method. Items that were in the same cluster in both the *iclust* and Ward's solutions were standardised and averaged into a single variable for subsequent analyses. This step resulted in 220 observed variables (81 solo items and 139 clusters, henceforth "symptoms and syndromes"; see Table S1) that were carried forward in the primary analyses.

Data Analysis

All primary analyses were conducted according to the *a priori* analytic plan based on the methods in Forbes et al. (2021), using both hierarchical principal components analysis (i.e., an extended bass-ackward approach; Forbes, 2023b) and hierarchical clustering (i.e., Ward's hierarchical agglomerative clustering; Ward, 1963). The analytic pipeline is shown in Figure 1. Briefly, we randomly split the 14,762 participants into a primary sample ($n = 11,762$) and a hold-out sample ($n = 3,000$) to examine the robustness of the results. Both the primary and hold-out samples were analysed the same way, starting with a smoothed Spearman correlation matrix of the 220 observed variables in each sample. Matrices were smoothed using eigenvalue decomposition (Bock et al., 1988; Wothke, 1993) and the smoothed correlation matrices were nearly identical to the unsmoothed correlation matrices (Pearson's $r_s > .999$ and Spearman's $\rho_{hos} > .999$).

Higher order modeling

The first stage of analyses focused on points of convergence within each sample between the hierarchical principal components analysis (hPCA) approach and hierarchical clustering. We used oblique (oblimin rotation) principal components analysis rather than factor analysis to maximize computational economy given the complex hierarchy with many levels and components expected. The number of components extracted was guided by parallel analysis, but we required at least three variables to have a unique primary loading $\geq .4$ (i.e., with all cross-loadings $< .4$) on each dimension so future work can operationalize the constructs as latent variables. After redundant components were removed from each solution (correlations $> .9$, and congruence coefficients $> .95$ for all variables in a chain), the hierarchy clustering solution in each sample (Figures S2 and S3) was examined for convergence/divergence with the hPCA structure to determine which principal components in the latter were possible statistical artefacts (see also Forbes, 2023b). These consensus higher order structures were compared in detail between samples and methods as well as based on summary statistics (e.g., congruence coefficients for

⁴ See the supplementary materials for information on a deviation from the *a priori* analytic plan, lowering a planned threshold for *iclust* to allow sufficient dimension reduction.

principal components and Baker's Gamma for the two clustering dendrograms) and combined based on their points of agreement between methods and samples (see Figure 1).

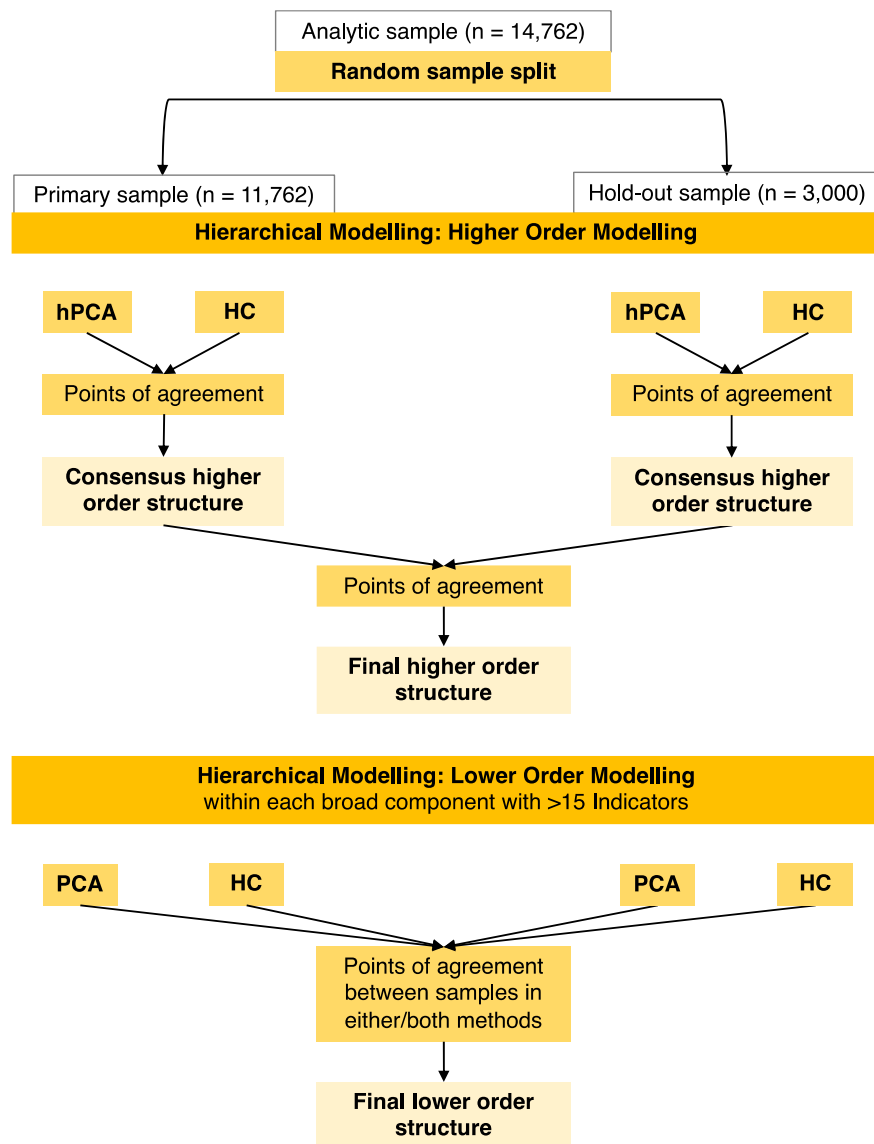


Figure 1. The analytic pipeline for the primary analyses.

hPCA = hierarchical principal components analyses using the extended bass-ackward approach; HC = Ward's hierarchical agglomerative clustering; PCA = principal components analysis.

Lower order modeling

Sizeable principal components at the bottom of the hierarchy (i.e., >15 variables with varied content) were analysed with another round of PCA to determine whether narrower interpretable components emerged within that set of variables. Item sets for each lower order PCA were identified by selecting symptoms that loaded $\geq |.3|$ on the principal components in both samples, had a loading $\geq |.3|$ in one sample mirrored by a cluster assignment in the other sample, or were assigned to the same construct in both hierarchical clustering solutions (see Figures S2 and S3). To examine robustness of the structures between methods, we compared the lower order PCA results to the lower levels of the hierarchical clustering solutions that were estimated in the higher

order modeling stage (Figures S2 and S3). The constructs that emerged in the lower order modeling were only interpreted as meaningful if they replicated between the primary and hold-out samples using either of the statistical approaches (i.e., based on the points of agreement between samples using either or both statistical methods).

For both stages of the hierarchical modeling, points of agreement between *samples* were defined as a variable with a component loading $\geq .3$ on the same construct and/or being assigned to the same cluster in both samples; points of agreement between *methods* were defined as a variable being assigned to the same construct using both principal components and hierarchical clustering analyses. The final higher order and lower order structures based on these points of agreement between samples and methods is the focus of interpretation here, but details of all code, output including the traditional back-ward solutions, all output for the higher order and lower order principal components analyses, and correlations among all principal components are available on the OSF page for the project (<https://osf.io/9v3gf/>).

Results

Higher Order Modelling

Primary sample

Parallel analysis indicated a maximum of 30 components to be extracted, but only the first six components had at least three variables with a unique primary loading $\geq .4$, so we extracted one to six components. After we removed the redundant components from the hPCA solution, we compared the pruned hPCA solution with the hierarchical clustering solution. When hPCA components were not evident in the hierarchical clustering solution, we removed them from the consensus higher-order structure (see supplementary materials for details on each artefact and Figure S4A for the consensus higher-order structure in the primary sample).

There was marked convergence in the higher order constructs that emerged in each method: *Internalizing*, *Externalizing*, *Thought Disorder*, *Neurodevelopmental and Cognitive Difficulties*, *Somatoform*, and *Mania/Low Detachment*—the latter as separate clusters that agglomerated in the upper levels of the clustering solution.⁵ Primary differences in the upper levels of the clustering solution versus the hPCA solution included an *Eating Pathology* cluster splitting from *Internalizing*, and a broader *Somatoform* cluster that incorporated many of the uncontrollable physical symptoms that were part of the *Thought Disorder* component (see dashed lines in Figure S2). The upper levels of the hPCA and hierarchical clustering were also relatively consistent in the combination of *Externalizing* with *Mania/Low Detachment*, and of *Internalizing* with *Somatoform* and *Eating Pathology*—the latter labelled *Emotional Dysfunction* to be consistent with the existing literature (e.g., Watson et al., 2022). There was mixed evidence on whether to include a single

⁵ To facilitate the readability of this text, throughout the Results and Discussion we use *Italicized Capitalized Labels* for the higher order constructs found in this study, and *italicized lowercase labels* for the lower order constructs. Construct names from the literature or from other studies are presented in plain text.

overarching dimension at the apex of the hierarchy (e.g., 82.7% of the variables loaded $> .3$ on the first unrotated principal component, clustering stopping rules⁶ indicated stopping agglomeration at 2 clusters); it was tentatively included to help resolve the placement of the *Neurodevelopmental and Cognitive Difficulties* component, which was related to all three of the superspectra. Because the superspectra varied in breadth across the samples and methods, they are outlined with dashed lines in Figure S4A to denote uncertainty about the nature of these constructs.

Hold-out sample

Parallel analysis indicated a maximum of 27 components to be extracted, but only the first seven components had at least three variables with a unique primary loading $\geq .4$, so we extracted one to seven components. As in the primary sample, both methods included *Internalizing*, *Externalizing*, *Thought Disorder*, *Neurodevelopmental and Cognitive Difficulties*, and *Mania/Low Detachment*—the latter again as separate clusters that agglomerated in the upper levels of the clustering solution. In contrast to the primary sample, *Eating Pathology* and *Harmful Substance Use* also emerged as distinct constructs across both methods. Further, a *Somatoform* component mirroring the findings in the primary sample was evident in the six-component hPCA solution as well as in the hierarchical clustering solution, so it was retained as a spectrum in the consensus higher order structure for the hold-out sample (Figure S4B). Primary differences in the upper levels of the clustering solution versus the hPCA solution included uncontrollable physical symptoms loading on the *Thought Disorder* component instead of clustering with *Somatoform*, and only three core detachment symptoms agglomerating with the mania cluster (see Figure S3). The upper levels of both solutions in the hold-out sample were the same as for the primary sample.

Final higher order structure

Although the exact patterns of component loadings and points of agglomeration in the clustering analyses were not the same between the samples (see supplementary materials for full details), there were clear patterns of consistency in the constructs that emerged. Where there was consistency in the assignment of variables to constructs, we included them in the final higher order structure (see Figure 2).

Lower Order Modelling: Points of agreement between samples

Six of the eight spectra at the bottom of the higher order structure had ≥ 15 variables with varied content that converged in their assignment on each construct between samples and/or methods. As we mentioned earlier, we further analysed these sets of variables with another round of PCA and hierarchical clustering (*lower order modeling*; see Figure 1) to explore whether narrower robust and interpretable constructs emerged. There was strong consistency between samples such that the mean correlation of the component loadings for variables that loaded $\geq .3$ on either instance of a component (i.e., in either sample) was $r = .90$ (range .58–1.00). Figure 2

⁶ Following Watts et al. (2021), we considered consensus across three stopping rules (i.e., Calinski-Harabasz, Davies-Bouldin, and Silhouette) to identify where to stop agglomerating clusters.

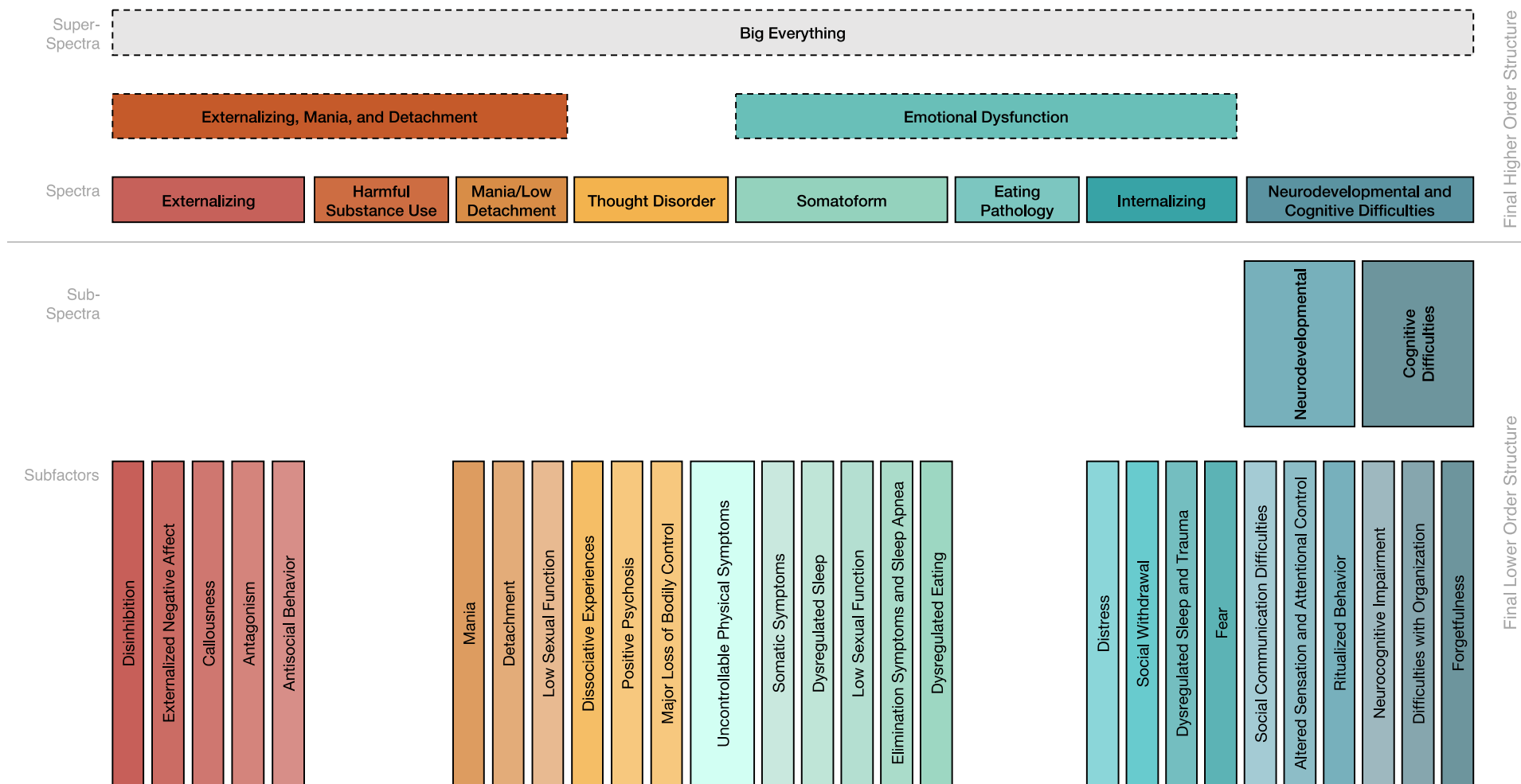


Figure 2. Summary model based on the points of agreement across methods (extended bass-ackwards and hierarchical clustering) and samples (primary and hold out). Constructs outlined with dashed edges emerged in all solutions but were less consistent in their content across the methods and samples. The ordering of the constructs from left to right does not carry systematic meaning.

shows the final lower order structure nested under the final higher order structure. Symptoms and syndromes that comprise each of these lower and higher order dimensions are shown in Figure 3, and the specific symptoms within each of those constructs are listed in Table S1. We now briefly summarize these lower order results.

Within *Externalizing* (Figure 3A), five robust lower order constructs emerged: *disinhibition*, *externalized negative affect*, *callousness*, *antagonism*, and *antisocial behavior*. All of these constructs except *antisocial behavior* were replicated across samples and methods; *antisocial behavior* was replicated in the two PCAs across samples but not in the hierarchical clustering. Within *Mania/Low Detachment* (Figure 3B), three robust lower order constructs replicated across samples and methods: *mania*, *detachment*, and *low sexual function*. Within *Thought Disorder* (Figure 3C), four robust lower order constructs emerged: *uncontrollable physical symptoms*, *dissociative experiences*, *positive psychosis*, and *major loss of bodily control*. All of these constructs except *uncontrollable physical symptoms* replicated across samples and methods; *uncontrollable physical symptoms* was replicated in the two PCAs across samples, but clustered with *Somatoform* in the hierarchical clustering solutions. An *uncontrollable physical symptoms* construct therefore also emerged under *Somatoform*, replicated across samples and methods⁷, and these two *uncontrollable physical symptoms* constructs were merged in the final structure (Figures 2 and 3C). Other lower order constructs that emerged under *Somatoform* (Figure 3D) included *somatic symptoms*, *dysregulated sleep*, *low sexual function*, *dysregulated eating*, and *elimination symptoms and sleep apnea*. All of these constructs were replicated across the two PCAs, but the variables loading on *dysregulated sleep*, *low sexual function*, and *dysregulated eating* in the PCAs also cross-loaded on and clustered with *Internalizing*, *Mania/Low Detachment*, and *Eating Pathology*, respectively. Within *Internalizing* (Figure 3E), four robust lower order constructs replicated across samples and methods: *distress*, *social withdrawal*, *dysregulated sleep and trauma*, and *fear*.

Finally, the lower order modeling in the *Neurodevelopmental and Cognitive Difficulties* spectrum initially only converged on two broad *Neurodevelopmental* and *Cognitive Difficulties* domains that each still had ≥ 15 varied indicators, so these subsets of variables underwent a third round of PCA to examine whether robust lower order constructs were evident. We found three lower order constructs in each subspectrum that replicated across samples and methods. For *Neurodevelopmental* (Figure 3F): *social communication difficulties*, *altered sensation and attentional control*, and *ritualized behavior*. For *Cognitive Difficulties* (Figure 3G): *neurocognitive impairment*, *difficulties with organization*, and *forgetfulness*.

Symptoms that did not load on any lower order constructs, or that loaded inconsistently between samples and methods, were not included in the final consensus structure. In total,

⁷It was possible for the *uncontrollable physical symptoms* construct to be replicated in the PCAs under both *Thought Disorder* and *Somatoform* because most indicators for *uncontrollable physical symptoms* cross-loaded on both spectra in the higher order analyses.

A

Spectra	Externalizing				
Subfactors	Disinhibition	Externalized Negative Affect	Callousness	Antagonism	Antisocial Behavior
Symptoms and Syndromes	Reckless risk-taking <u>Dyscontrol</u> ³ Hazardous substance use Responsibility avoidance Uncharacteristic behavior ³ Inappropriate hyperactivity Inappropriate behavior Inappropriate sexual behavior Oppositionality Carelessness Low effort Feigned injury or illness Kleptomanic tension Anti-authoritarianism	Irritability Frustration Angry outbursts Argumentativeness Resentment Suspiciousness <u>Dyscontrol</u> ³ <u>Blame externalization</u>	Reduced empathy Lack of empathy Lack of remorse Emotional indifference Callous disregard <u>Grandiose disparagement</u> <u>Grandiose entitlement</u>	Attention seeking Manipulative skill Manipulative deceitfulness Manipulativeness <u>Grandiose entitlement</u> Cheating for advantage <u>Factitious deceitfulness</u> <u>Blame externalization</u> Using others' things Deliberately annoying behavior Deceitful irresponsibility <u>Grandiose disparagement</u>	Deliberately harmful behavior Physical aggression <u>(low) Factitious deceitfulness</u>

B

Spectra	Harmful Substance Use	Mania/Low Detachment		
Subfactors		Mania	Detachment	Low Sexual Function
Symptoms and Syndromes	Substance use disorder symptoms Illicit and prescription substance use Tobacco use Hazardous substance use Cannabis use Alcohol use	Euphoric activation Attention seeking Euphoria Expansive mood expression <u>Excitement</u> Unusually absorbed planning Eroticism Socially adaptive communication	Restricted emotional expression Emotional indifference Reclusiveness Depressed mood and anhedonia ³ Unemotionality <u>Romantic avoidance</u> <u>(low) Excitement</u>	Sexual anhedonia Low sexual interest Sexual arousal and orgasmic dysfunction <u>Romantic avoidance</u>

C

Spectra	Thought Disorder			Somatoform
Subfactors	Dissociative Experiences	Positive Psychosis	Major Loss of Bodily Control	Uncontrollable Physical Symptoms
Symptoms and Syndromes	Dissociative dyscontrol Episodic amnesia Dissociation Dissociative amnesia Catatonia Sensory numbing Suicide attempt Cognitive distortions of trauma Multiple personalities Time distortion <u>Choking sensations</u> Night terrors <u>Unusual experiences</u> <u>Uncharacteristic behavior</u> ²	Delusions of love and grandeur Thought insertion and deletion Delusions Ritualized greeting Unusual social approach <u>Eccentricity</u> ⁴ Synesthesia Unusual experiences <u>Uncharacteristic behavior</u> ³ Perceptual dysregulation Repetitive rituals and compulsions Decreased need for sleep Twilight confusion	Cataplexy Seizures Waxy flexibility	Twitching Tics Dystonias Tremors Nystagmus Paresthesia <u>Goosebumps/chills</u> Blurred vision Restless legs* Muscular rigidity Speech impediments* Mutism* <u>Choking sensations</u> <u>Sleepwalking/sleepwalking*</u> <u>Eccentricity</u> ^{4*} <u>Psychomotor impairment</u> ^{3*} Selective mutism*

D

Spectra	Somatoform				
Subfactors	Somatic Symptoms	Dysregulated Sleep	Low Sexual Function	Dysregulated Eating	Elimination Symptoms and Sleep Apnea
Symptoms and Syndromes	Gastrointestinal distress Constipation <u>Bloating</u> Headaches Health anxiety Frequent healthcare seeking Physical panic Aches and pains <u>Psychomotor impairment</u> ³ Hot flushes <u>Goosebumps/chills</u>	Insomnia Early sleep and awakening Nightmares Sleepiness	Sexual anhedonia Sexual arousal and orgasmic dysfunction	Overeating Weight gain Night eating <u>Bloating</u>	Enuresis Encopresis Sleep apnea

E

Spectra	Eating Pathology	Internalizing			
Subfactors		Distress	Social Withdrawal	Dysregulated Sleep and Trauma	Fear
Symptoms and Syndromes	Binge-related distress Weight and shape concerns Overeating Dysmorphic appearance concerns ³ Weight gain Rapid eating Restricted food intake	Irritability Resentment Frustration Emotional lability Dyscontrol ³ <u>Depressed mood and anhedonia</u> ³ Self-derogation Suicidality Suspiciousness Guilt/shame proneness Fear of rejection and abandonment Dissociation Morning depression Psychomotor agitation Intrusive thoughts <u>Dysmorphic appearance concerns</u> ³ Psychological panic Anxiousness	Low social engagement ³ Reclusiveness Agoraphobia and social fears Increased introversion Home separation anxiety <u>Depressed mood and anhedonia</u> ³	Insomnia Early sleep and awakening Nightmares Night terrors Traumatic intrusions and avoidance Grief	Separation anxiety Fear of dying Sleep separation anxiety Somatic separation anxiety Fear of rejection and abandonment Specific phobia <u>Dysmorphic appearance concerns</u> ³ Anxiety tolerance Health anxiety Agoraphobia and social fears Hypervigilance and situational avoidance Home separation anxiety

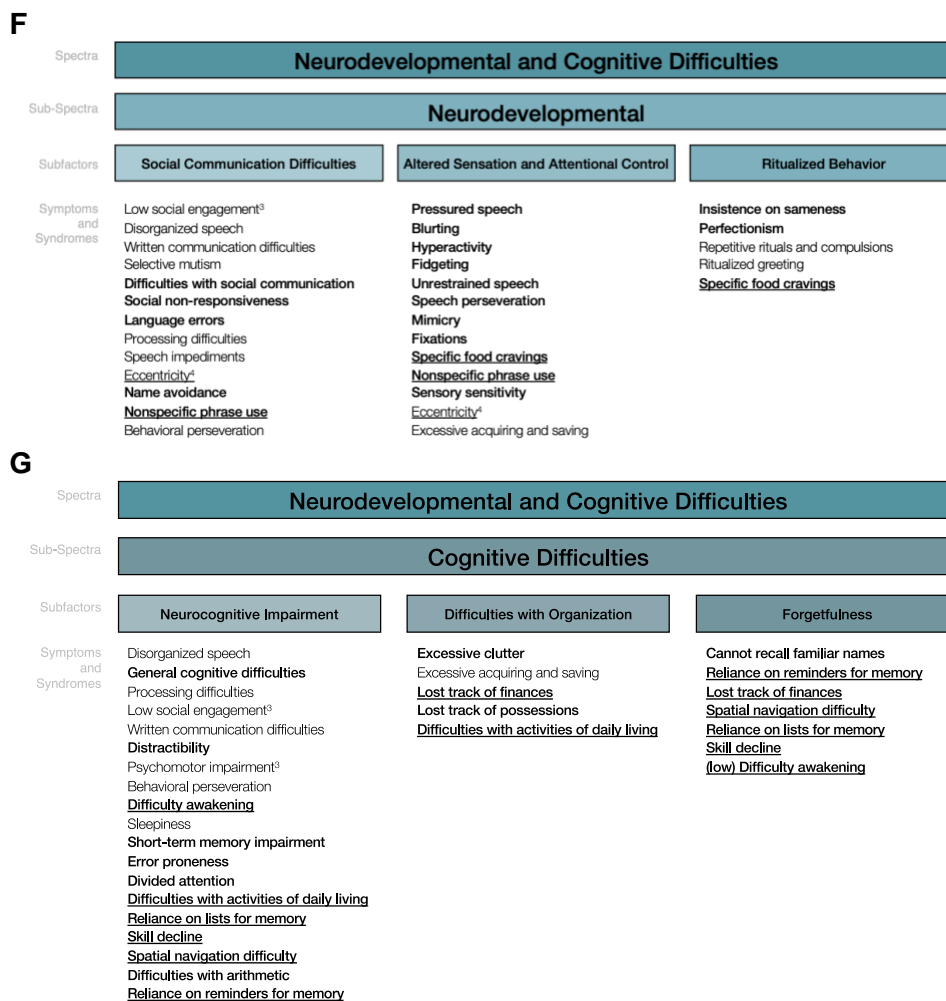


Figure 3. Symptoms and syndromes of the lower order constructs (i.e., *subfactors*) in the hierarchy. *Note.* Symptoms included in each construct at the level of ‘Symptoms and Syndromes’ are listed in Table S1. The order of the constructs listed under each subfactor is based on the strength of the component loadings in the lower order PCAs in the primary sample, with the exceptions of *Harmful Substance Use* and *Eating Pathology*, which are sorted based on the strength of the component loadings in the hold-out sample where they emerged as robust constructs, and *uncontrollable physical symptoms*. The two *uncontrollable physical symptoms* subfactors described in text were merged, and the variables under this construct are sorted based on their average loadings across the two components in the main sample. Indicators with an asterisk * were part of one of the *uncontrollable physical symptoms* subfactors, but not both. All component loadings are provided on the OSF page for the project (<https://osf.io/9v3qf/>).

Bolded constructs are specific to one spectrum (or sub-spectrum in the case of *Neurodevelopmental* and *Cognitive Difficulties* dimensions).

Underlined constructs cross-loaded within the spectrum (or sub-spectrum in the case of *Neurodevelopmental* and *Cognitive Difficulties* dimensions).

Thin text constructs cross-load in another spectrum.

Greyed text indicates when the placement of a construct was only replicated in the hierarchical clustering solutions across samples, not in the hierarchical principal components analyses.

Superscript numbers ³ and ⁴ indicate where a construct cross-loaded in three or four places, respectively; all others cross-load in two places.

186 (84.5%) of the 220 symptom and syndrome constructs were included (see Table S1 for symptoms and syndromes not assigned in the final structure). Most of the symptoms and syndromes were assigned to a single construct ($n = 130$, 69.9%), with the remainder cross-loading on two ($n = 49$, 26.3%), three ($n = 6$, 3.2%), or four ($n = 1$, 0.5%) spectra or subspectra; constructs that are unique to a single spectrum are bolded in Figure 3. The final structure included 27 subfactors, 2 subspectra⁸, 8 spectra, and 3 tentative superspectra. We note that the assignment of constructs to levels of the hierarchy is ambiguous in some cases (e.g., spectra with no nested lower order constructs might be better described as subspectra, and the current subspectra of *Neurodevelopmental* and *Cognitive Difficulties* might be better described as spectra nested under a superspectrum). Regardless of this uncertainty, we have used these labels to aid clarity in the description and discussion of each level of the hierarchy shown in Figures 2-3.

Discussion

This study reorganized *DSM-5* symptoms into an empirically derived hierarchical framework of clinical phenomena. Here, we focus on two key features of our findings. Firstly, whether the structure departs from the quantitative psychopathology literature when taking a symptom-level approach to analyses with a broader construct pool and without a reliance on *DSM*-defined diagnoses. Secondly, whether exemplar *DSM-5* disorders were supported in our structure to the extent that their symptoms formed coherent constructs or syndromes. Overall, our symptom level structure (Figures 2-3) had noteworthy convergence with the quantitative psychopathology literature—particularly with the six core spectra of the HiTOP model (Figure S1) and the Alternative Model of Personality Disorders in Section III of the *DSM-5* (see Figure S6). Further, in nearly all cases⁹ there was evident heterogeneity in *DSM*-defined disorders such that their symptoms spanned multiple syndromes, subfactors, and/or (sub-)spectra (Figure 4).

At the top of the hierarchy, we called the single overarching dimension the *Big Everything* (i.e., spanning personality pathology, other psychopathology, and symptoms of cognitive functioning; Littlefield et al., 2021), departing from the convention of a ‘general psychopathology’ or ‘p-factor’ label to avoid reifying the common factor found here, which was extracted by default given the methods we used. The *Emotional Dysfunction* superspectrum mirrored the existing quantitative psychopathology literature (e.g., Watson et al., 2022), but the superspectrum made up of *Externalizing*, *Harmful Substance Use*, and *Mania/Low Detachment* departed from the distinct externalizing and psychosis superspectra we might have expected based on the extant literature

⁸From an hPCA perspective, there was some evidence of three other intermediary subspectra in the lower order analyses: (1) a dimension made up of *distress* and *social withdrawal* under the *Internalizing* spectrum; (2) a dimension made up of *dysregulated sleep*, *elimination symptoms*, and *low sexual function* under the *Physical Symptoms* spectrum; and (3) a dimension made up of *altered sensation and attentional control* with *ritualized behavior* under the *Neurodevelopmental* sub-spectrum. Although we did not focus on this level of the hierarchy here, we note them briefly in the event that they are useful for future studies.

⁹ Some disorders such as insomnia, delayed ejaculation, erectile disorder, and female orgasmic disorder had all of their symptoms nested within a single syndrome—albeit mixed with symptoms of other diagnoses.

(e.g., Kotov et al., 2020; Krueger et al., 2021). As for the *Big Everything*, we consider these two superspectra as tentative constructs because there were inconsistencies between samples and methods in their breadth. Below the superspectra were eight other higher order constructs that were robust between samples and methods: *Externalizing*, *Harmful Substance Use*, *Mania/Low Detachment*, *Thought Disorder*, *Somatoform*, *Eating Pathology*, *Internalizing*, and *Neurodevelopmental and Cognitive Difficulties*. We now briefly discuss each one compared to the quantitative psychopathology literature and the *DSM*.

Externalizing

The content of the *Externalizing* spectrum was largely consistent with the existing personality pathology literature (e.g., Sleep et al., 2021). It expanded on the HiTOP externalizing dimensions (disinhibited externalizing, antagonistic externalizing, and antisocial behavior), with *callousness* and *externalized negative affect* forming separate subfactors, resulting in a somewhat narrower *antagonism* subfactor than in the HiTOP model. Features of grandiosity loaded on both *callousness* and *antagonism*, and together with the content captured by *externalized negative affect* reflect the antagonistic externalizing spectrum in HiTOP (Krueger et al., 2021). Corresponding to these differences in structure, the prominent disinhibited externalizing and antagonistic externalizing spectra from the HiTOP model were demoted in the structure to form two of the five subfactors under a broad externalizing spectrum (Figure 2). Harmful substance use was notably distinct from externalizing, in contrast to its placement in HiTOP, as discussed further in the next section.

Consistent with the *DSM* constructs that map onto the externalizing spectrum in HiTOP, symptoms of intermittent explosive disorder, oppositional defiant disorder, and conduct disorder all mapped cleanly onto *Externalizing* (Figure 4). By contrast, most ADHD symptoms that are part of disinhibited externalizing in HiTOP (e.g., hyperactivity, distractibility, and [low] perfectionism) were largely related to the *Neurodevelopmental* subspectrum here (Figure 4; see also Michelini et al., 2019).

Harmful Substance Use

The *Harmful Substance Use* spectrum was remarkably internally consistent, with all symptoms derived from the substance use disorders falling here (Figure S5) and only hazardous substance use cross-loading elsewhere (i.e., on *disinhibition* under *Externalizing*). In fact, the symptoms from the substance use disorders were so closely related that all but hazardous use formed a single syndrome at the first stage of dimension reduction. This finding is likely due in part to 1) the conditional dependence among the substance use disorder symptoms, which all refer specifically to the use of substances (i.e., logically, endorsement of each symptom required that the individual had used a substance in the past 12 months), and 2) the use of only 13 items to assess substance use disorder symptoms (see Watts et al., 2021 for discussion of how crude measurement can contribute to observed unidimensionality).

The fact that the hazardous-use criterion both remained separate from the other symptoms during the clustering phase and cross-loaded under the *Externalizing* spectrum mirrors Watts et al.'s (2023d) finding that the alcohol use disorder criterion of recurrent use of alcohol in hazardous situations had a uniquely strong relationship with externalizing, compared to other alcohol use disorder criteria. That the remaining substance use disorder symptoms formed a spectrum distinct from *Externalizing* adds to a growing body of evidence that substance use and many features of substance use disorder are likely not as closely related to externalizing as the HiTOP model implies (Watts et al., 2023d).

Instead, the consequences of substance use, such as role interference and continuing to drink despite social/interpersonal harm appear most closely linked with externalizing because those symptoms incorporate or subsume functional impairment and reflect the deficits in behavioral and emotional control captured in externalizing (e.g., Martin et al., 2014; Watts et al., 2023c). This suggests that the placement of harmful substance use under externalizing in HiTOP is likely to be a consequence of using *DSM* diagnoses of substance use disorders as the units of analysis in quantitative psychopathology research (Watts & Boness, 2022). Its separation from externalizing may also reflect that harmful substance use is etiologically multifactorial and not linked simply to externalizing psychopathology, but also to negative affectivity (e.g., Boness et al., 2021).

Mania/Low Detachment

In the HiTOP model, mania cross-loads between thought disorder and internalizing, and detachment represents a distinct spectrum unrelated to mania. By comparison, the *Mania/Low Detachment* spectrum found here differed in both content and placement to HiTOP, but was consistent with other quantitative psychopathology research that has found mania and detachment to form a superordinate dimension (e.g., Stanton et al., 2021b). Figure 4 shows how the symptoms of a *DSM*-defined manic episode spread across components of our hierarchical structure, suggesting that modeling these symptoms as a single coherent dimension may not be supported empirically. Instead, it seems that the core mania symptoms of elevated and expansive mood, unusually absorbed planning, and increased energy may represent an empirically supported syndrome while recklessness, hyperactive cognition and behaviors, and psychosis related to mania are scattered elsewhere throughout the structure. However, it is important to note that all symptoms were assessed in a 12-month recall period, which excludes the hallmark episodicity of mania that has been found to be a distinguishing characteristic in structural models of psychopathology (Jonas et al., 2023) and may be capturing trait levels of high energy and positive affect. This finding highlights the importance of more quantitative psychopathology research tailored to assessing the episodicity of mania.

The *detachment* subfactor included some domains that are consistent with those of HiTOP (e.g., anhedonia, depressivity, intimacy avoidance, and reclusiveness), but there were also a few exceptions. First, suspiciousness loaded on *externalized negative affect* and *distress*, rather than

detachment (see also Forbes et al., 2021). Second, *low sexual function* emerged from *detachment* in the lower order analyses, which is in contrast to the literature informing HiTOP (see Forbes et al., 2017 for a review) but has strong face validity with sexual anhedonia, low sexual interest, and romantic avoidance closely mirroring features of detachment. As discussed below, *low sexual function* also loaded on *Somatoform*—perhaps capturing the equifinality of sexual function relating to both psychological (e.g., cognitive and emotional processes, reward processing) and physiological (e.g., hormonal, endocrine, vascular) mechanisms (e.g., Basson et al., 2000). Notably, *detachment* did not load together with *Thought Disorder* to form a psychosis superspectrum (cf. Kotov et al., 2020).

Thought Disorder

The *Thought Disorder* spectrum stood alone in the structure here, rather than relating to *mania*, or loading together with *detachment* to form a psychosis superspectrum (cf. Kotov et al., 2021). The traits that compose thought disorder in HiTOP (e.g., eccentricity, cognitive and perceptual dysregulation, and unusual beliefs and experiences) were largely preserved in the *positive psychosis* subfactor. By contrast, the components of HiTOP thought disorder (psychotic, disorganized, inexpressivity, avolition)—derived from symptom-level work on the dimensional structure of psychosis (e.g., Kotov et al., 2016; Longenecker et al., 2022)—did not cohere, mirroring the heterogeneity of schizophrenia shown in Figure 4. For example, although positive symptoms loaded under *Thought Disorder*, symptoms mapping onto the HiTOP construct of disorganized behavior loaded on the *Neurodevelopment and Cognitive Difficulties* spectrum, avolition loaded on the *detachment* subfactor, and inexpressivity-related symptoms spanned both *Neurodevelopment and Cognitive Difficulties* as well as *Mania/Low Detachment*. These patterns mirror work that has found positive psychosis symptoms to load on psychoticism whereas other symptoms load elsewhere in the structure of normal personality (Cicero et al., 2019), as well as findings that reality distortion symptoms represent a unique core of psychosis whereas disorganized behavior and negative symptoms are associated with cognitive impairment (e.g., Cowan et al., 2023; Dibben et al., 2009; Dominguez et al., 2009; Ventura et al., 2010).

The *Thought Disorder* spectrum found in these analyses also included additional subfactors that are not included in HiTOP, namely *dissociative experiences* and *major loss of bodily control*. These findings are consistent with the close association of dissociation and catatonia with psychosis (e.g., Kotov et al., 2020; Longden et al., 2020). Finally, the *uncontrollable physical symptoms* subfactor resembled the *DSM-5's* functional neurologic disorder/conversion disorder (e.g., symptoms of dystonias, tremors, paresthesia, blurred vision, changes in speech, and choking sensations; APA, 2013). That this dimension spanned the *Thought Disorder* and *Somatoform* spectra is also consistent with the primary associated features of functional neurologic disorder being both dissociative symptoms (e.g., depersonalization, derealization, and dissociative amnesia) and somatic symptoms (e.g., pain and fatigue; APA, 2013).

Somatoform

Our *Somatoform* spectrum was broader than in past research (see Watson et al., 2020), including not only the *uncontrollable physical symptoms* subfactor mentioned above but also the hallmark symptoms of somatic symptom disorder and illness anxiety under *somatic symptoms*, as well as *elimination symptoms and sleep apnea, dysregulated sleep, low sexual function, and dysregulated eating*. As mentioned in the results, symptoms of these latter three domains were nonspecific to *Somatoform*, cross-loading under *Internalizing, Mania/Low Detachment, and Eating Pathology*, respectively. Overall, the *Somatoform* spectrum had a shared core with the somatoform domain in HiTOP, but also pulled across additional physical symptoms from a variety of disorders (e.g., from mood and anxiety disorders, see Figure S5).

Eating Pathology

The *Eating Pathology* spectrum included nearly all symptoms of *DSM-5* eating disorders that were part of the final model¹⁰ (i.e., anorexia nervosa, bulimia nervosa, binge eating disorder; Figure 4), mirroring the three *DSM* constructs that anchor the eating pathology domain in HiTOP. The cognitive, emotional, and behavioral symptoms captured in the bingeing, restricted eating, and weight and shape concerns syndromes were all specific to the *Eating Pathology* spectrum and cover many of the core symptom domains of eating pathology (e.g., Forbush et al., 2013). However, some of the *Eating Pathology* content cross-loaded on other constructs (e.g., overeating and weight gain were a part of *dysregulated eating* under *Somatoform*; dysmorphic appearance concerns were part of the *fear* and *distress* domains under *Internalizing*). Finding the *Eating Pathology* spectrum to be distinct from the *Internalizing* spectrum departs from much of the quantitative psychopathology literature and HiTOP, perhaps because of its increased representation in these symptom analyses. Even so, more research is needed to clarify whether eating pathology and internalizing are intertwined (e.g., Forbush & Watson, 2013; Forbush et al., 2018) or more loosely related at the level of an emotional dysfunction superspectrum, as was found here (e.g., Watson et al., 2022).

Internalizing

The *Internalizing* spectrum included the core content domains expected based on the literature underpinning HiTOP, spanning *distress, social withdrawal, dysregulated sleep and trauma*, and *fear* subfactors. However, it did not include the more peripheral domains of *sexual problems, eating pathology, and mania* that are included under internalizing in HiTOP. The *distress* subfactor represents the theoretical core of *Internalizing* (i.e., negative affect; Clark et al., 1994; Watson & Clark, 1984), differentiated from *externalized negative affect* under *Externalizing* by the

¹⁰Eating disorder symptoms were overrepresented in the symptoms that did *not* appear in the final model. Appetite loss, low body weight, pathological laxative use, weight loss, and purging did not load consistently, so were not included in the final structure. Notably, these symptoms did not load together with the other eating pathology symptoms in *any* of the solutions between methods or samples. Need for tube feeding or nutritional supplements due to malnutrition, and excessive exercise also had such low endorsement (i.e., tube feeding) or associations with other symptoms (i.e., supplements and exercise) that they were dropped prior to the main analyses.

presence of indicators such as emotional lability, depressed mood and anhedonia, suicidality, and guilt and shame proneness (vs. angry outbursts, argumentativeness, and blame externalization as unique indicators for *externalized negative affect*). The *social withdrawal* subfactor (new to internalizing vs. HiTOP) encompassed behavioral aspects of social disengagement; the *dysregulated sleep and trauma* subfactor included nearly all trauma-specific symptoms, as well as grief and sleep disturbances, including nightmares; and the *fear* subfactor reflected phobic aspects of anxiety. Notably, although the *DSM* phobic anxiety disorders and separation anxiety disorder mapped cleanly onto the *Internalizing* spectrum here, other prototypical internalizing disorders such as MDD, GAD, and PTSD disintegrated (see Figure 4), which corresponded to marked dispersion of the symptom components listed under the distress subfactor in the HiTOP model (Figure S1).

Neurodevelopmental and Cognitive Difficulties

Finally, the *Neurodevelopmental and Cognitive Difficulties* spectrum split into separate (1) *Neurodevelopmental* and (2) *Cognitive Difficulties* subspectra. The *Neurodevelopmental* subspectrum aligns with prior work that indicates close associations among symptoms of autism, ADHD, and OCD rituals and compulsions (e.g., Kushki et al., 2019; Rommelse et al., 2011), as well as other findings that inattention and social/communication difficulties combine in structural models of psychopathology (e.g., Michelini et al., 2019; Stanton et al., 2021a, 2021b). The *Cognitive Difficulties* subspectrum is novel and could provide a path to incorporating neurocognitive disorder symptoms—such as difficulties in complex attention, executive function, and perceptual-motor functioning—into HiTOP, strengthening the case for a cognitive dysfunction dimension of psychopathology (e.g., Abramovitch et al., 2021). Together, these two subspectra encompassed a variety of symptoms that are currently housed elsewhere in the HiTOP model (e.g., cognitive symptoms from HiTOP's somatoform spectrum; perseveration and restricted affect from internalizing; some symptoms of inexpressivity from thought disorder; and distractibility and [low] perfectionism from externalizing). It is likely that providing a larger variety of potential indicators—relative to most previous quantitative psychopathology research—allowed these broad dimensions to form, but more work is needed to establish their validity and utility.

One noteworthy feature of the lower order constructs composing the *Neurodevelopmental* and *Cognitive Difficulties* domains was the overlap in symptoms and syndromes that loaded on both *social communication difficulties* and *neurocognitive impairment*, potentially reflecting equifinality of these symptoms. For example, symptoms of disorganized speech, processing difficulties, low social engagement, written communication difficulties, and behavioral perseveration included: “I had difficulty finding the words I wanted to say,” “I had difficulty understanding the meaning of what I was reading,” “I did not initiate conversations with others,” “It was hard to convey my thoughts in writing,” and “I had trouble changing how I was doing something even when it wasn't working,” respectively—all of which could capture either social communication difficulties or cognitive impairment, depending on the underlying causes.

Summary

Taken together, these results diverge from the structure of the current HiTOP framework in a number of interesting ways: 1) the addition of neurodevelopmental and cognitive difficulties domains; 2) the demotion and restructuring of externalizing (i.e., superspectrum → spectrum; spectra → subfactors); 3) harmful substance use forming an independent spectrum from externalizing; 4) eating pathology forming an independent spectrum from internalizing; 5) the reshuffling of sexual problems, mania, and detachment; 6) the addition of a number of subfactors under externalizing, thought disorder, somatoform, and internalizing; and 7) some reorganization of the lower levels of the hierarchy with cross-loadings that appeared likely to represent cases of equifinality, and separate constructs forming for physical, emotional, behavioural, and cognitive symptoms. With further testing of these findings (e.g., using the new HiTOP self-report and interview measures; Simms et al., 2021) and the accumulation of sufficient evidence, some of these points of difference between our structure and HiTOP could be incorporated in a formal revision of the HiTOP model (Forbes et al., 2023c). This seems to be a more likely path for some findings (e.g., incorporating neurodevelopmental and cognitive difficulties domains) than others (e.g., making changes to the placement of mania based on these data, which lack assessment of episodicity in symptoms).

Relative to the breadth of the *DSM*, the scope of our hierarchical structure expands substantially on the scope of HiTOP—from approximately 71 *DSM* disorders in the current HiTOP model (see Forbes, 2023a) to 167 disorders in our hierarchical structure. Although there are noteworthy absences from our hierarchical structure¹¹, there are also important additions, particularly in the representation of dissociative, elimination, sleep-wake, trauma-related, neurodevelopmental, and neurocognitive disorders. This expansion of coverage of psychopathology did not engender a large loss in parsimony compared to HiTOP because there is considerable repetition of symptoms across putatively distinct *DSM* disorders (i.e., the additional 96 disorders only added 178 additional symptoms; see Forbes et al., 2023b). By contrast to the extensive symptom repetition in the structure of the *DSM*, only seven symptom domains repeated three or more times in our structure—most of which highlighted the known nonspecificity of negative affect (Stanton et al., 2023).

As mentioned earlier, our results also indicated heterogeneity in most of the *DSM*-defined constructs, and this was most pronounced for diagnoses with the deepest historical roots (e.g., schizophrenia and MDD). It is not reasonable to propose a total reorganization of the symptoms and disorders described in the *DSM* or *ICD* on the basis of this study alone, but one tractable implication for the revisions of these traditional classification systems could be incorporating

¹¹For example, from feeding and eating disorders: rumination, pica, and core symptoms domains of purging and appetite/weight loss were not included; from OCRDs: hair pulling, skin picking, and body-focused repetitive behaviors; from disruptive, impulse-control, and conduct disorders: pyromania and kleptomania. The chapters of paraphilic disorders and gender dysphoria were also not incorporated at all.

something like the meta-structure proposed by Andrews et al. (2009), which included broad domains that parallel the ones we found here: neurocognitive (called *Cognitive Difficulties* here), neurodevelopmental, psychosis (*Thought Disorder*), emotional (*Internalizing*), and externalizing. This could directly incorporate some of the benefits of a hierarchical structure without requiring a total reorganization.

Limitations and Future Directions

Arguably, the data used here are better suited for examining the finer-grained structure of psychopathology—such as homogeneous syndromes—than analyses of existing self-report measures developed to assess predetermined constructs (often defined by the *DSM*). By starting with the full symptom pool described in the *DSM-5* and using fully randomized item presentation without skip outs, we were able to follow the empirical patterns in the data to derive the constructs found here. However, our approach also generated several particularly important limitations to consider in interpreting these results. First, to make measurement of all *DSM-5* symptoms tractable, we decontextualized symptoms to isolate their core content, which will have erased the differences between some symptoms that are apparent when they are considered in context (e.g., insomnia while withdrawing from a substance versus insomnia due to worrying; Saunders, 2021; Zachar, 2023). Second, all symptoms were assessed using the same 12-month time scale despite the fact that time scale sometimes delineates the patterns of interest among symptoms for certain phenomena (e.g., manic episodes, binge eating episodes, panic attacks); the 12-month time scale also blurs the distinction between symptoms and traits (DeYoung et al., 2022). Third, the specific framing of the self-reported symptoms will have shaped individuals' responses and may not have captured culture-specific experiences of some phenomena, so future research could examine variations on the assessment approach taken here (Saunders, 2021).

Another specific issue to consider in this non-exhaustive list of limitations is the absence of a 'not applicable' response option. Participants were instructed to select '*Not at all true (Never)*' for items that did not apply to their experiences, but this approach will have conflated the mechanisms of response to some items (Waller, 1989). Specifically, there were 44 items in the initial item pool (6.5%) that could potentially be inapplicable to someone's experiences (e.g., experiences specifically referring to trauma, gambling, substance use, or sexual activity), which likely attenuated the true inter-item correlations for these constructs with other items, potentially explaining the exclusion of some of these items from the structure (see Table S1). On the other hand, the absence of a not applicable option will also have inflated the correlations *within* this subset of items (Waller, 1989), which likely contributed to the formation of tightly bound gambling, trauma, and substance use syndromes. Future research should consider including a not applicable option where relevant.

More broadly, the translation of the *DSM-5* into a self-report measure does not do justice to the intended use of the symptom criteria described in the manual, including the loss of special status for core diagnostic symptoms (e.g., worry in GAD; depressed mood and anhedonia in MDD;

APA, 2013). Focusing on self-reported symptoms made the scope of the study possible, but this study design also precluded the assessment of signs only observable by others, and relied on individuals' insight into their symptoms, giving authority to the first-person perspective (Ward & Clack, 2019). Future work should go further in centering lived experiences in the refinement of the items, naming of the resulting constructs, and understanding the acceptability and utility of the hierarchical and dimensional approach to assessment and diagnosis. It will also be essential to understand which aspects of these results—particularly the fine-grained levels of the structure—are robust to other approaches to measurement (i.e., using alternative measures, time frames, multi-method or multi-informant approaches, as well as within-person assessment) and across intersectional conceptualizations of identity (e.g., in a variety of sociodemographic and culturally and linguistically diverse samples). A guiding principle for future work should be iteration between refining the internal validity (e.g., factor structure, psychometric reliability) and external validity (e.g., prediction of other variables) of hierarchical structures of psychopathology (Forbes et al., 2023c).

Conclusion

This study reorganized the symptoms described in the *DSM-5* into a data-driven model based on the patterns in individuals' self-reported experiences of the symptoms, aiming to overcome some of the prominent limitations of both the *DSM-5* and HiTOP. The final structure proposed here represents an important step towards a comprehensive, empirically derived and supported classification system for psychopathology. It provides a preliminary map of homogeneous clinical phenotypes that could advance research and practice beyond the *DSM* by offering target constructs for neuroscience, genetics, and clinical psychology research as well as research in other frameworks moving away from traditional classification systems (e.g., RDoC, dynamic systems, process-based therapy, clinical staging; see Eaton et al., 2023; Forbes et al., 2023a; Rief et al., in press). Ultimately, the hope is that improving the reliability and validity of the constructs we study will provide new insights on risk factors and mechanisms in psychopathology, leading to improved treatment outcomes.

References

- Abramovitch, A., Short, T., & Schweiger, A. (2021). The C Factor: Cognitive dysfunction as a transdiagnostic dimension in psychopathology. *Clinical Psychology Review*, 86, 102007. <https://doi.org/10.1016/j.cpr.2021.102007>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). <https://doi.org/10.1176/appi.books.9780890425787>
- Anderson, J., Snider, S., Sellbom, M., Krueger, R., & Hopwood, C. (2014). A comparison of the DSM-5 Section II and Section III personality disorder structures. *Psychiatry Research*, 216(3), 363–372. <https://doi.org/10.1016/j.psychres.2014.01.007>
- Andrews, G., Goldberg, D. P., Krueger, R. F., Carpenter, W. T., Hyman, S. E., Sachdev, P., & Pine, D. S. (2009). Exploring the feasibility of a meta-structure for DSM-V and ICD-11: Could it improve utility and validity? *Psychological Medicine*, 39(12), 1993–2000. <https://doi.org/10.1017/S0033291709990250>
- Basson, R., Berman, J., Burnett, A., Derogatis, L., Ferguson, D., Fourcroy, J., ... & Whipple, B. (2000). Report of the international consensus development conference on female sexual dysfunction: definitions and classifications. *The Journal of Urology*, 163(3), 888–893. [https://doi.org/10.1016/S0022-5347\(05\)67828-7](https://doi.org/10.1016/S0022-5347(05)67828-7)
- Bock, R. D., Gibbons, R., & Muraki, E. (1988). Full-Information Item Factor Analysis. *Applied Psychological Measurement*, 12(3), 261–280. <https://doi.org/10.1177/014662168801200305>
- Boness, C. L., Watts, A. L., Moeller, K. N., & Sher, K. J. (2021). The etiologic, theory-based, ontogenetic hierarchical framework of alcohol use disorder: A translational systematic review of reviews. *Psychological Bulletin*, 147(10), 1075–1123. <https://doi.org/10.1037/bul0000333>
- Cicero, D. C., Jonas, K. G., Li, K., Perlman, G., & Kotov, R. (2019). Common Taxonomy of Traits and Symptoms: Linking Schizophrenia Symptoms, Schizotypy, and Normal Personality. *Schizophrenia Bulletin*, 45(6), 1336–1348. <https://doi.org/10.1093/schbul/sbz005>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412–1427. <https://doi.org/10.1037/pas0000626>
- Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology*, 103(1), 103–116. <https://doi.org/10.1037/0021-843X.103.1.103>

- 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: A Journal of the Association for Psychological Science*, 14(3), 419–436. <https://doi.org/10.1177/1745691618810696>
- Cowan, H. R., Williams, T. F., Schiffman, J., Ellman, L. M., & Mittal, V. A. (2023). Mapping Psychosis Risk States Onto the Hierarchical Taxonomy of Psychopathology Using Hierarchical Symptom Dimensions. *Clinical Psychological Science*. <https://doi.org/10.1177/21677026221146178>
- Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, 11(1), 126. <https://doi.org/10.1186/1741-7015-11-126>
- Derks, E. M., Thorp, J. G., & Gerring, Z. F. (2022). Ten challenges for clinical translation in psychiatric genetics. *Nature Genetics*, 1–9. <https://doi.org/10.1038/s41588-022-01174-0>
- DeYoung, C. G., Chmielewski, M., Clark, L. A., Condon, D. M., Kotov, R., Krueger, R. F., Lynam, D. R., Markon, K. E., Miller, J. D., Mullins-Sweatt, S. N., Samuel, D. B., Sellbom, M., South, S. C., Thomas, K. M., Watson, D., Watts, A. L., Widiger, T. A., Wright, A. G. C., & the HiTOP Normal Personality Workgroup. (2022). The distinction between symptoms and traits in the Hierarchical Taxonomy of Psychopathology (HiTOP). *Journal of Personality*, 90(1), 20–33. <https://doi.org/10.1111/jopy.12593>
- Dibben, C. R., Rice, C., Laws, K., & McKenna, P. J. (2009). Is executive impairment associated with schizophrenic syndromes? A meta-analysis. *Psychological Medicine*, 39(3), 381–392. <https://doi.org/10.1017/S0033291708003887>
- Dominguez, M. de G., Viechtbauer, W., Simons, C. J., van Os, J., & Krabbendam, L. (2009). Are psychotic psychopathology and neurocognition orthogonal? A systematic review of their associations. *Psychological Bulletin*, 135(1), 157–171. <https://doi.org/10.1037/a0014415>
- Eaton, N. R., Bringmann, L. F., Elmer, T., Fried, E. I., Forbes, M. K., Greene, A. L., ... & Waszczuk, M. A. (2023). A review of approaches and models in psychopathology conceptualization research. *Nature Reviews Psychology*, 2(10), 622–636.
- First, M. B., Reed, G. M., et al. (in prep). Major Proposals for Changes to DSM/ICD Classification Systems: Rationale, Past Actions, and Future Prospects.
- Forbes, M. K. (2023a). Implications of the Symptom-Level Overlap Among DSM Diagnoses for Dimensions of Psychopathology. *Journal of Emotion and Psychopathology*, 1(1), Article 1. <https://doi.org/10.55913/joep.v1i1.6>
- Forbes, M. K. (2023b). Improving hierarchical models of individual differences: An extension of Goldberg's bass-ackward method. *Psychological Methods*. <https://doi.org/10.1037/met0000546>

- Forbes, M. K., Baillie, A. J., Eaton, N. R., & Krueger, R. F. (2017). A place for sexual dysfunctions in an empirical taxonomy of psychopathology. *The Journal of Sex Research*, 54(4-5), 465-485.
- Forbes, M. K., Fried, E. I., & Vaidyanathan, U. (2023a). Studying fine-grained elements of psychopathology to advance mental health science. *Journal of Psychopathology and Clinical Science*, 132(7), 793.
- Forbes, M. K., Neo, B., Nezami, O. M., Fried, E. I., Faure, K., Michelsen, B., Twose, M., & Dras, M. (2023b). Elemental psychopathology: Distilling constituent symptoms and patterns of repetition in the diagnostic criteria of the DSM-5. *Psychological Medicine*.
<https://doi.org/10.1017/s0033291723002544>
- Forbes, M. K., Ringwald, W. R., Allen, T., Cicero, D. C., Clark, L. A., DeYoung, C. G., ... Wright, A. G. (2023c). Principles and procedures for revising the Hierarchical Taxonomy of Psychopathology. <https://doi.org/10.31234/osf.io/xr48p>
- Forbes, M. K., Sunderland, M., Rapee, R. M., Batterham, P. J., Calear, A. L., Carragher, N., Ruggero, C., Zimmerman, M., Baillie, A. J., Lynch, S. J., Mewton, L., Slade, T., & Krueger, R. F. (2021). A Detailed Hierarchical Model of Psychopathology: From Individual Symptoms up to the General Factor of Psychopathology. *Clinical Psychological Science*, 9(2), 139–168. <https://doi.org/10.1177/2167702620954799>
- Forbush, K. T., Chen, P. Y., Hagan, K. E., Chapa, D. A., Gould, S. R., Eaton, N. R., & Krueger, R. F. (2018). A new approach to eating-disorder classification: Using empirical methods to delineate diagnostic dimensions and inform care. *International Journal of Eating Disorders*, 51(7), 710-721. <https://doi.org/10.1002/eat.22891>
- Forbush, K. T., & Watson, D. (2013). The structure of common and uncommon mental disorders. *Psychological Medicine*, 43(1), 97-108. <https://doi.org/10.1017/s0033291712001092>
- Forbush, K. T., Wildes, J. E., Pollack, L. O., Dunbar, D., Luo, J., Patterson, K., Petruzzi, L., Pollpeter, M., Miller, H., Stone, A., Bright, A., & Watson, D. (2013). Development and validation of the Eating Pathology Symptoms Inventory (EPSI). *Psychological Assessment*, 25(3), 859–878. <https://doi.org/10.1037/a0032639>
- Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, 172, 96–102. <https://doi.org/10.1016/j.jad.2014.10.010>
- Hofmann, S. G., & Hayes, S. C. (2019). The Future of Intervention Science: Process-Based Therapy. *Clinical Psychological Science*, 7(1), 37–50.
<https://doi.org/10.1177/2167702618772296>
- Jonas, K., Busso, C., Lian, W., Docherty, A. R., Shabalin, A. A., Carlson, G., Ruggero, C., Bromet, E., & Kotov, R. (2023). Where is Mania in the Hierarchical Taxonomy of Psychopathology: Internalizing, Thought Disorder, or Novel Spectrum?
<https://doi.org/10.31234/osf.io/nrgav>

- Kotov, R., Foti, D., Li, K., Bromet, E. J., Hajcak, G., & Ruggero, C. J. (2016). Validating Dimensions of Psychosis Symptomatology: Neural Correlates and 20-year Outcomes. *Journal of Abnormal Psychology, 125*(8), 1103–1119. <https://doi.org/10.1037/abn0000188>
- Kotov, R., Jonas, K. G., Carpenter, W. T., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Hobbs, K., Reininghaus, U., Slade, T., South, S. C., Sunderland, M., Waszczuk, M. A., Widiger, T. A., Wright, A. G. C., Zald, D. H., Krueger, R. F., Watson, D., & Workgroup, H. U. (2020). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry, 19*(2), 151–172. <https://doi.org/10.1002/wps.20730>
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. 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. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454–477. <https://doi.org/10.1037/abn0000258>
- Kotov, R., Krueger, R. F., Watson, D., Cicero, D. C., Conway, C. C., DeYoung, C. G., Eaton, N. R., Forbes, M. K., Hallquist, M. N., Latzman, R. D., Mullins-Sweatt, S. N., Ruggero, C. J., Simms, L. J., Waldman, I. D., Waszczuk, M. A., & Wright, A. G. C. (2021). The Hierarchical Taxonomy of Psychopathology (HiTOP): A Quantitative Nosology Based on Consensus of Evidence. *Annual Review of Clinical Psychology, 17*(1), 83–108. <https://doi.org/10.1146/annurev-clinpsy-081219-093304>
- Kotov, R., Perlman, G., Gámez, W., & Watson, D. (2015). The structure and short-term stability of the emotional disorders: A dimensional approach. *Psychological Medicine, 45*(8), 1687–1698. <https://doi.org/10.1017/S0033291714002815>
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879–1890. <https://doi.org/10.1017/S0033291711002674>
- Krueger, R. F., Hobbs, K. A., Conway, C. C., Dick, D. M., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Keyes, K. M., Latzman, R. D., Michelini, G., Patrick, C. J., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Tackett, J., Waldman, I., Waszczuk, M. A., ... Workgroup, H. U. (2021). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry, 20*(2), 171–193. <https://doi.org/10.1002/wps.20844>
- Krueger, R. F., & Markon, K. E. (2014). The Role of the DSM-5 Personality Trait Model in Moving Toward a Quantitative and Empirically Based Approach to Classifying Personality and Psychopathology. *Annual Review of Clinical Psychology, 10*(1), 477–501. <https://doi.org/10.1146/annurev-clinpsy-032813-153732>

- Kushki, A., Anagnostou, E., Hammill, C., Duez, P., Brian, J., Iaboni, A., Schachar, R., Crosbie, J., Arnold, P., & Lerch, J. P. (2019). Examining overlap and homogeneity in ASD, ADHD, and OCD: A data-driven, diagnosis-agnostic approach. *Translational Psychiatry*, 9, 318. <https://doi.org/10.1038/s41398-019-0631-2>
- Levin-Aspenson, H., Watson, D., Clark, L. A., & Zimmerman, M. (2021). What is the general factor of psychopathology? Consistency of the p factor across samples. *Assessment*, 28(4), 1035-1049. <https://doi.org/10.1177/1073191120954921>
- Levin-Aspenson, H. F. (2023). To fully leverage fine-grained clinical phenomena, we have to think beyond DSM based concepts and the presumption of diagnostic kinds. *Journal of Psychopathology and Clinical Science*. <https://doi.org/10.1037/abn0000876>
- Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. *Annual Review of Clinical Psychology*, 12, 435–463. <https://doi.org/10.1146/annurev-clinpsy-021815-093122>
- Longden, E., Branitsky, A., Moskowitz, A., Berry, K., Bucci, S., & Varese, F. (2020). The relationship between dissociation and symptoms of psychosis: A meta-analysis. *Schizophrenia Bulletin*, 46(5), 1104–1113. Advance online publication. <https://doi.org/10.1093/schbul/sbaa037>
- Longenecker, J. M., Haas, G. L., & Salisbury, D. F. (2022). Hierarchical Symptom Components in Early Psychosis. *Schizophrenia Bulletin*, 48(4), 893–901. <https://doi.org/10.1093/schbul/sbac048>
- Maples, J. L., Carter, N. T., Few, L. R., Crego, C., Gore, W. L., Samuel, D. B., Williamson, R. L., Lynam, D. R., Widiger, T. A., Markon, K. E., Krueger, R. F., & Miller, J. D. (2015). Testing whether the DSM-5 personality disorder trait model can be measured with a reduced set of items: An item response theory investigation of the Personality Inventory for DSM-5. *Psychological Assessment*, 27(4), 1195–1210. <https://doi.org/10.1037/pas0000120>
- Martin, C. S., Langenbucher, J. W., Chung, T., & Sher, K. J. (2014). Truth or consequences in the diagnosis of substance use disorders. *Addiction*, 109(11), 1773-1778. <https://doi.org/10.1111/add.12615>
- McGorry, P. D., Hickie, I. B., Yung, A. R., Pantelis, C., & Jackson, H. J. (2006). Clinical staging of psychiatric disorders: A heuristic framework for choosing earlier, safer and more effective interventions. *The Australian and New Zealand Journal of Psychiatry*, 40(8), 616–622. <https://doi.org/10.1080/j.1440-1614.2006.01860.x>
- Michelin, G., Barch, D. M., Tian, Y., Watson, D., Klein, D. N., & Kotov, R. (2019). Delineating and validating higher-order dimensions of psychopathology in the Adolescent Brain Cognitive Development (ABCD) study. *Translational Psychiatry*, 9(1), Article 1. <https://doi.org/10.1038/s41398-019-0593-4>

- Newson, J. J., Pastukh, V., & Thiagarajan, T. C. (2021). Poor Separation of Clinical Symptom Profiles by DSM-5 Disorder Criteria. *Frontiers in Psychiatry*, 12, 775762. <https://doi.org/10.3389/fpsyt.2021.775762>
- 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(1), 59–70. <https://doi.org/10.1176/appi.ajp.2012.12070999>
- Revelle W. (1979). Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behavioral Research*, 14(1), 57–74. https://doi.org/10.1207/s15327906mbr1401_4
- Rief, W., Hofmann, S. G., Berg, M., Forbes, M. K., Pizzagalli, D. A., Zimmerman, J., Fried, E., & Reed, G. M. (in press). Do we need a novel framework for classifying psychopathology? A discussion paper. *Clinical Psychology in Europe*
- Ringwald, W. R., Forbes, M. K., & Wright, A. G. C. (2023). Meta-analysis of structural evidence for the Hierarchical Taxonomy of Psychopathology (HiTOP) model. *Psychological Medicine*, 53(2), 533–546. <https://doi.org/10.1017/S0033291721001902>
- Rommelse, N. N. J., Geurts, H. M., Franke, B., Buitelaar, J. K., & Hartman, C. A. (2011). A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience & Biobehavioral Reviews*, 35(6), 1363–1396. <https://doi.org/10.1016/j.neubiorev.2011.02.015>
- Saunders, J. B. (2021). Symptoms and Their Relationship to Diagnosis: A Commentary on “Symptom Descriptions in Psychopathology: How Well Are They Working for Us?” (Wilshire et al., 2021). *Clinical Psychological Science*, 9(3), 343–346. <https://doi.org/10.1177/2167702621990356>
- Simms, L. J., Wright, A. G., Cicero, D., Kotov, R., Mullins-Sweatt, S. N., Sellbom, M., ... & Zimmermann, J. (2022). Development of measures for the Hierarchical Taxonomy of Psychopathology (HiTOP): A collaborative scale development project. *Assessment*, 29(1), 3-16.
- Sleep, C. E., Crowe, M. L., Carter, N. T., Lynam, D. R., & Miller, J. D. (2021). Uncovering the structure of antagonism. *Personality Disorders: Theory, Research, and Treatment*, 12(4), 300–311. <https://doi.org/10.1037/per0000416>
- Stanton, K. (2020). Increasing diagnostic emphasis on negative affective dysfunction: Potentially negative consequences for psychiatric classification and diagnosis. *Clinical Psychological Science*, 8(3), 584–589. <https://doi.org/10.1177/2167702620906147>
- Stanton, K., Balzen, K. M., DeFluri, C., Brock, P., Levin-Aspenson, H. F., & Zimmerman, M. (2023). Negative mood dysregulation loads strongly onto common factors with many forms of psychopathology: Considerations for assessing nonspecific symptoms. *Assessment*. <https://doi.org/10.1177/10731911231174471>

- Stanton, K., DeLucia, E. A., Brown, M. F. D., & McDonnell, C. G. (2021 a). Advancing understanding of the classification of broad autism phenotype and attention-deficit/hyperactivity disorder symptom dimensions within the Hierarchical Taxonomy of Psychopathology. *Personality and Mental Health*, 15(2), 113–123.
<https://doi.org/10.1002/pmh.1498>
- Stanton, K., Khoo, S., McDonnell, C. G., Villalongo Andino, M., Sturgeon, T., & Aasen, L. (2021b). An Initial Investigation of the Joint Classification of Hypomania- and Neurodevelopmental Disorder-Relevant Dimensions Within the Hierarchical Taxonomy of Psychopathology. *Assessment*, 10731911211055670.
<https://doi.org/10.1177/10731911211055670>
- Tiego, J., Martin, E. A., DeYoung, C. G., Hagan, K., Cooper, S. E., Pasion, R., Satchell, L., Shackman, A. J., Bellgrove, M. A., & Fornito, A. (2023). Precision behavioral phenotyping as a strategy for uncovering the biological correlates of psychopathology. *Nature Mental Health*, 1(5), Article 5. <https://doi.org/10.1038/s44220-023-00057-5>
- Üstün, B., & Kennedy, C. (2009). What is “functional impairment”? Disentangling disability from clinical significance. *World Psychiatry*, 8(2), 82–85. <https://doi.org/10.1002/j.2051-5545.2009.tb00219.x>
- Ventura, J., Thames, A. D., Wood, R. C., Guzik, L. H., & Helleman, G. S. (2010). Disorganization and reality distortion in schizophrenia: A meta-analysis of the relationship between positive symptoms and neurocognitive deficits. *Schizophrenia Research*, 121(1), 1–14. <https://doi.org/10.1016/j.schres.2010.05.033>
- Waller, N. G. (1989). The Effect of Inapplicable Item Responses on the Structure of Behavioral Checklist Data: A Cautionary Note. *Multivariate Behavioral Research*, 24(1), 125–134.
https://doi.org/10.1207/s15327906mbr2401_8
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236–244.
<https://doi.org/10.1080/01621459.1963.10500845>
- Ward, T., & Clack, S. (2019). From symptoms of psychopathology to the explanation of clinical phenomena. *New Ideas in Psychology*, 54, 40-49.
<https://doi.org/10.1016/j.newideapsych.2019.01.004>
- Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96(3), 465–490. <https://doi.org/10.1037/0033-2909.96.3.465>
- Watson, D., Levin-Aspenson, H. F., Waszczuk, M. A., Conway, C. C., Dalgleish, T., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Hobbs, K. A., Michelini, G., Nelson, B. D., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Waldman, I., Witthöft, M., Wright, A. G. C., ... Workgroup, H. U. (2022). Validity and utility of Hierarchical Taxonomy of

- Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry*, 21(1), 26–54. <https://doi.org/10.1002/wps.20943>
- Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., Stasik, S. M., & Ruggero, C. J. (2012). Development and Validation of New Anxiety and Bipolar Symptom Scales for an Expanded Version of the IDAS (the IDAS-II). *Assessment*, 19(4), 399–420. <https://doi.org/10.1177/1073191112449857>
- Watts, A. L., & Boness, C. L. (2022). A proposal to revise the title of HiTOP's substance abuse subconstruct. *Open Science Framework*, 1-10. <https://osf.io/zmscu/>
- Watts, A. L., Boness, C. L., Loeffelman, J. E., Steinley, D., & Sher, K. J. (2021). Does crude measurement contribute to observed unidimensionality of psychological constructs? A demonstration with DSM-5 alcohol use disorder. *Journal of Abnormal Psychology*, 130(5), 512–524. <https://doi.org/10.1037/abn0000678>
- Watts, A. L., Greene, A. L., Bonifay, W., & Fried, E. I. (2023a). *A critical evaluation of the p-factor literature*. PsyArXiv. <https://doi.org/10.31234/osf.io/7yrnp>
- Watts, A. L., Latzman, R. D., Boness, C. L., Kotov, R., Keyser-Marcus, L., DeYoung, C. G., ... & Ramey, T. (2023b). New approaches to deep phenotyping in addictions. *Psychology of Addictive Behaviors*, 37(3), 361.
- Watts, A. L., Sher, K. J., Heath, A. C., Steinley, D., & Brusco, M. (2023c). 'General addiction liability' revisited. <https://doi.org/10.17605/OSF.IO/XDSNP>
- Watts, A. L., Watson, D., Heath, A. C., & Sher, K. J. (2023d). Alcohol use disorder criteria exhibit different comorbidity patterns. *Addiction*, 118, 1457-1468. <https://doi.org/10.1111/add.16121>
- Williams, T. F., Scalco, M. D., & Simms, L. J. (2018). The construct validity of general and specific dimensions of personality pathology. *Psychological Medicine*, 48(5), 834–848. <https://doi.org/10.1017/S0033291717002227>
- Wills, G. B., & Lessler, J. T. (1999). *Question Appraisal System*. <https://docplayer.net/147787-Question-appraisal-system-qas-99.html>
- Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In *Kenneth A. Bollen and J. Scott Long (Editors), Testing structural equation models*. Sage Publications.
- Zachar, P. (2023). Non-specific psychopathology: A once and future concept. *World Psychiatry*, 22(1), 154–155. <https://doi.org/10.1002/wps.21043>