The Development of Preliminary HiTOP Internalizing Spectrum Scales

David Watson
University of Notre Dame

Miriam K. Forbes
Macquarie University

Holly F. Levin-Aspenson
University of Notre Dame

Camilo J. Ruggero
University of North Texas

Yuliya Kotelnikova
University of New Orleans

Shereen Khoo
University of Notre Dame

R. Michael Bagby
University of Toronto

Matthew Sunderland
University of Sydney

Praveetha Patalay
University College London

Roman Kotov
Stony Brook University

The reported research was approved by the Institutional Review Boards at the University of Notre Dame, the University of North Texas, and the University of Toronto Scarborough. Correspondence concerning this paper should be addressed to David Watson, Department of Psychology, 390 Corbett Family Hall, University of Notre Dame, Notre Dame, IN, 46556, USA. Email: db.watson@nd.edu.

Abstract

As part of a broader project to create a comprehensive self-report measure for the Hierarchical Taxonomy of Psychopathology (HiTOP) consortium, we developed preliminary scales to assess internalizing symptoms. The item pool was created in four steps: (a) clarifying the range of content to be assessed; (b) identifying target constructs to guide item writing; (c) developing formal definitions for each construct; and (d) writing multiple items for each construct. This yielded 430 items assessing 57 target constructs. Responses from a heterogeneous scale development sample (N = 1,870) were subjected to item-level factor analyses based on polychoric correlations. This resulted in 39 scales representing a total of 213 items. The psychometric properties of these scales replicated well across the development sample and an independent validation sample (N = 496 adults). Internal consistency analyses established that most scales assess relatively narrow forms of psychopathology. Structural analyses demonstrated the presence of a strong general factor. Additional analyses of the 35 non-sexual dysfunction scales revealed a replicable four-factor structure with subdimensions we labeled Distress, Fear, Body Dysmorphia, and Mania. A final set of analyses established that the internalizing scales varied widely—and consistently—in the strength of their associations with neuroticism and extraversion.

KEYWORDS: internalizing psychopathology, mania, obsessive-compulsive and related disorders, sexual dysfunction, scale construction, factor analysis, discriminant validity

The Development of Preliminary HiTOP Internalizing Spectrum Scales

The Hierarchical Taxonomy of Psychopathology (HiTOP) is an international consortium that is using data from structural studies to construct a quantitative nosological system (Kotov et al., 2017). Assessment is a crucial part of this effort. HiTOP-consistent measures are needed in the clinic to provide practitioners with a viable alternative to assessment methods based on traditional diagnostic systems (Ruggero et al., 2019). HiTOP-congruent measures also are needed in structural research to clarify the classification of conditions whose placement currently is inconsistent or unclear (e.g., mania, borderline pathology).

To facilitate future structural work, the group concluded that it was necessary to create a new, comprehensive self-report measure. Self-report scales have been developed to measure virtually all forms of psychopathology, and these instruments potentially could be combined to create a reasonably comprehensive assessment battery. However, they employ a diverse array of instructions, time frames (e.g., past two weeks, past month, past 12 months) and response formats (e.g., agree-disagree, frequency-based, intensity-based). These different methods complicate structural research by creating artifactual method effects (i.e., all other things being equal, variables assessed using the same format and time frame will be more strongly related than those based on different methods; see Johnson, Rosen, & Djurdjevic, 2011; Reio, 2010). Consequently, it is beneficial to assess all types of psychopathology using the same method.

Given the size and complexity of this task, the HiTOP Measurement Workgroup decided to construct this measure in multiple stages. In Phase 1 of the project, five different subgroups were created to develop preliminary sets of scales falling within the major spectra identified in structural research (Kotov et al., 2017); specifically, these five subgroups were charged with the assessment of (a) internalizing, (b) disinhibited and antagonistic externalizing, (c) thought

disorder, (d) detachment, and (e) somatoform psychopathology. Each group was given substantial autonomy in creating preliminary Phase 1 scales, although efforts were made to minimize the overlap between them. In Phase 2 of the project, the preliminary scales from these five groups will be subjected to joint analyses to begin the process of creating a comprehensive measure.

This paper documents the Phase 1 activities of the Internalizing Group. The term *internalizing* originated in the child psychopathology literature to describe forms of psychopathology that involve "problems within the self" (Achenbach, 1966, p. 10), as opposed to maladaptive interactions with the environment. Structural researchers subsequently adopted this term to describe a spectrum of symptoms and disorders characterized primarily by emotional dysfunction and/or related forms of behavioral avoidance (Watson et al., in press). The internalizing spectrum consistently emerges in structural studies of psychopathology (Watson et al., in press); it consists of several subfactors, including distress (e.g., major depression, dysthymic disorder, generalized anxiety disorder [GAD], posttraumatic stress disorder [PTSD]), fear (e.g., panic disorder, agoraphobia, social phobia, specific phobia), eating pathology (e.g., bulimia nervosa, anorexia nervosa, binge eating disorder), and sexual problems (e.g., low sexual desire, sex-related distress, difficulties with arousal). Finally, mania is interstitial, showing important connections to both internalizing and thought disorder (Watson et al., in press).

Development of the Item Pool

Establishing the Scope of the Domain

The item pool was developed in four stages. The initial stage (conducted in September/October 2017) focused on clarifying the scope of the assessment project. We used Figure 2 (which displays disorders) and Figure 3 (which lists symptom components and maladaptive traits) in Kotov et al. (2017) as the starting point for identifying the range of content

to be subsumed within internalizing. Based on these sources, we established that the measure should assess content related to the depressive disorders, the anxiety disorders (including symptoms related to GAD, social phobia, panic disorder, agoraphobia, separation anxiety disorder, and specific phobia), the trauma-related disorders (especially PTSD), obsessive-compulsive disorder [OCD], sexual problems, irritability, and the maladaptive traits shown in Figure 3 of Kotov et al. (2017) (e.g., hostility, perseveration, separation insecurity). As noted earlier, structural studies have indicated that eating pathology also clearly falls within the internalizing spectrum (Forbush et al., 2010; Forbush & Watson, 2013; Kotov et al., 2017). For purely practical reasons, however, eating pathology was assigned to the Somatoform Group due to its members having greater relevant expertise. Consequently, we will not be discussing eating pathology in this paper. To ensure continuity of efforts between the two groups, the Chair of the Internalizing Group (DW) was included as a member of the Somatoform Group.

Three other considerations merit some discussion. First, as noted previously, mania is an interstitial construct between internalizing and thought disorder in the HiTOP model (Watson et al., in press). The Measurement Workgroup decided that the Internalizing Group would assume primary responsibility for assessing mania. However, some Thought Disorder Group members contributed to the writing of mania items. The Thought Disorder Group also collected its own data on these items and participated in analyses of this interstitial content.

Second, sleep-wake disorders also show important links to internalizing (e.g., Koffel et al., 2016; Koffel & Watson, 2009; Watson et al., 2012). The HiTOP model of internalizing already includes content related to insomnia and lassitude/fatigue, which help to anchor two major dimensions of sleep pathology (Koffel, 2011). To expand our coverage of sleep pathology, we decided to assess nightmares as well. Nightmares have been linked to indicators of stress,

anxiety, panic, depression, and suicidality; however, they exhibit particularly strong and consistent associations with PTSD (e.g., Koffel, 2011; Koffel, Khawaja, & Germain, 2016; Levin & Nielsen, 2007; Watson, Stasik, Ellickson-Larew, & Stanton, 2015).

Third, the obsessive-compulsive and related disorders (OCRDs) are a new diagnostic class in *DSM-5* (American Psychiatric Association, 2013). OCD is the only OCRD currently included in the HiTOP model. The other OCRDs—hoarding disorder, body dysmorphic disorder, trichotillomania/hair-pulling disorder, and excoriation/skin-picking disorder—have not been examined in major structural studies of psychopathology, so their placement currently is unclear. Because OCD already was included in internalizing, we decided to assess content related to the other OCRDs as well. Based on recent factor analytic work (Watson, Stasik-O'Brien, Ellickson-Larew, & Stanton, 2018), we included five specific OCRD constructs in our item pool: body dissatisfaction, body preoccupation, excoriation, hoarding, and trichotillomania.

Identifying the Target Constructs

In the second stage (conducted in October/November 2017), we identified the specific target constructs to be assessed in our measure. In this phase of the process, we opted for splitting (i.e., modeling narrowly defined symptom dimensions), as opposed to lumping (i.e., focusing on broader constructs). Consistent with recommended psychometric guidelines (Clark & Watson, 2019; Watson, 2012), our goal here was to include multiple markers to define all of the symptom dimensions that potentially could emerge in subsequent structural analyses.

Again, the figures provided in Kotov et al. (2017) provided a starting point for compiling these constructs. These HiTOP figures were supplemented by consulting the structural/ assessment evidence establishing the existence of specific symptom dimensions within depression (e.g., Dornbach-Bender et al., 2017; Waszczuk, Kotov, Ruggero, Gaméz, & Watson,

2017; Watson et al., 2012), GAD (e.g., Mahoney et al., 2016), PTSD (e.g., Dornbach-Bender et al., 2017; Gootzeit, Markon, & Watson, 2015; Waszczuk et al., 2017; Watson et al., 2012; Yufik & Simms, 2010), OCD (e.g., Dornbach-Bender et al., 2017; Foa et al., 2002; Waszczuk et al., 2017; Watson et al., 2012; Watson & Wu, 2005), social anxiety (e.g., Dornbach-Bender et al., 2017; Waszczuk et al., 2017; Watson et al., 2012), panic/agoraphobia (e.g., Dornbach-Bender et al., 2017; Waszczuk et al. 2017), specific phobia (e.g., Cutshall & Watson, 2004; Dornbach-Bender et al., 2017; Waszczuk et al., 2017; Watson et al., 2012), mania (e.g., Dornbach-Bender et al., 2017; Ruggero et al. 2014; Waszczuk et al., 2017; Watson et al., 2012), and sexual problems (e.g., Forbes, Baillie, & Schniering, 2016a, 2016b). Finally, individual Group members were encouraged to suggest additional constructs in their areas of expertise.

This process eventually produced a set of 57 target constructs to be modeled in the initial item pool. These 57 constructs are shown in Table 1, along with a sample item for each one (the item writing process is described later). Table 1 also classifies these constructs into general areas of content (e.g., anxiety, sleep, depression). It should be emphasized that this classification scheme was introduced purely for the sake of convenience (most notably, to facilitate locating constructs within the model) and has no broader significance. For example, in addition to being a sleep-wake disorder in *DSM-5* (American Psychiatric Association, 2013), insomnia also is a symptom of several other disorders—including major depression, GAD, and PTSD—and could have been placed in any one of several different categories.

Writing Construct Definitions

The third stage (conducted in December 2017/January 2018) involved an iterative process of writing brief formal definitions for each target construct. During this process, Group members refined definitions falling within their areas of expertise. These construct definitions

subsequently served as a guide for item writing. Appendix A (included in the supplemental materials) provides the final definitions for all 57 constructs. Following the typical practice in internalizing symptom assessment (e.g., Watson et al. 2012), these constructs were conceptualized as unipolar in nature, such that all items were written in the keyed direction. In this regard, it should be noted that manifestations of low positive affect ("Anhedonia/low well-being") and high positive affect ("Well-Being/high positive affect") were modeled as two separate unipolar constructs.

Creating the Item Pool

The fourth stage was the formal creation of the item pool. This stage can be broken down into two main phases. Item writing was the focus of the first phase, which was conducted from January to May 2018. Group members were provided with the instructions and response format to be used in the measure, and were asked to "write as many non-redundant items as you can" for those constructs falling within their areas of expertise. Six Group members (MF, RK, HFLA, CR, MS, DW) contributed items at this stage. This item writing process yielded a total of 1,110 items. Supplemental Table S1 presents the number of items written for each target construct in the preliminary item pool. The number of items varied considerably across constructs, ranging from 8 (Sexual Pain) to 38 (Irritability/Hostility), with a mean of 19.5 items. We will refer to these 57 sets of rationally organized items as *homogenous item composites,* or HICs (see Clark & Watson, 2019; Hogan, 1983; Watson et al., 2012).

This preliminary pool was unwieldy and contained a large number of clearly redundant items. Consequently, the second stage (conducted in June/July 2018) involved selecting the best items for each HIC. Group members were asked to read through each HIC and to indicate the items that were the best indicators of the target construct, taking into account such issues as (a)

ensuring comprehensiveness of coverage and (b) item redundancy/overlap. Raters were allowed to nominate as many—or as few—items as they saw fit, and could select a maximum of five items as representing the best overall indicators of the construct.

To facilitate this rating task, the preliminary item pool was divided into three spreadsheets, each containing roughly 300-400 items. The first set included the mania, maladaptive trait, and sexual dysfunction HICs (401 items overall); six Group members (MF [sex items only], RK, HFLA, CR, MS, DW) rated these items. The second set included the sleep, depression, and PTSD HICs (310 items overall); five Group members (MF, RK, HFLA, PP, DW) evaluated these items. The final group included the anxiety, OCD, and OCRD HICs (399 items overall); five Group members (MF, RK, HFLA, MS, DW) rated these items.

For each target construct, we selected the highest rated items for inclusion in the final item pool. This yielded a final pool of 430 items. Table S1 shows the number of items contained in each reduced HIC in the final item pool. The number of retained items ranged from 6 to 11, with a mean of 7.5 items per HIC. Appendix A presents all of the retained items for each HIC.

Participants and Procedure

Development Sample

Assessment protocol. We collected data in the Development sample between September 2018 and May 2019. The assessment protocol was divided into four sections. The first section asked the respondents to provide basic demographic information, including gender, age, and race/ethnicity. To assess their clinical status, participants also were asked if they were "currently receiving psychological counseling/therapy for mental health issues" or had received "psychological counseling/therapy for mental health issues in the past."

The second section contained the 30-item short form of the Big Five Inventory-2 (BFI-2-S;

Soto & John, 2017). The BFI-2-S assesses the domains and facets of the five-factor model (FFM) of personality. Each BFI-2-S domain includes three two-item facet scales. The items are rated on a 5-point scale ranging from *disagree strongly* to *agree strongly.* We report results on the BFI-2-S Negative Emotionality and Extraversion domain scores.

The last two sections included the 430 items in the final item pool. Participants were asked whether there had been "significant times during the last 12 months during which the following statements applied to you"; they responded using a 4-point scale (*not at all, a little, moderately, a lot*). The internalizing items were split into two sections because of the skip-structure in the sexual dysfunction items; these items therefore were separated from the others. The third section of the protocol contained all 395 non-sexual dysfunction items, plus four "instructed response" validity checks (Meade & Craig, 2012) that were designed to make sure respondents were paying attention (e.g., "Please select 'Moderately' for this question"); the items were presented in a random order with items from the different HICs interspersed together.

The fourth section included the 35 sexual dysfunction items, plus three screening questions asking about sexual activity during the past year. All participants were presented with an initial set of 15 questions that assessed their level of interest in sex, as well as thoughts and emotions about sexual activity. The subsequent blocks of items were split up into experiences that required (1) sexual activity, including passionate kissing, foreplay, or masturbation (e.g., "I felt desire during sexual activity"; 12 items), (2) attempting to reach orgasm (e.g., "I was unable to reach orgasm"; 4 items), or (3) experiencing orgasm (e.g., "my orgasms were pleasurable"; 4 items). Each of these blocks began with a screening question asking whether or not the respondent had the relevant experience in the past 12 months. If they answered "No" to a screening question, they were instructed to skip to the end of the protocol.

Participants. We collected data from undergraduate students at the University of Notre Dame (N = 543), the University of North Texas (N = 782), and the University of Toronto Scarborough (N = 134). In addition, we used two different crowdsourcing platforms to recruit participants. First, we collected data from 328 adult participants via the recruiting service provided by Qualtrics. To be eligible for the study, interested participants completed an initial screening questionnaire indicating that they currently were receiving psychological treatment. Consequently, all of the Qualtrics participants currently were in psychological treatment (see Table 2). Second, we collected responses from 320 adult participants using the Prolific recruiting platform. To be eligible for the study, respondents indicated that they either were currently receiving psychological treatment or had received treatment in the past. Thus, the large majority of the Prolific participants had some history of psychological treatment. However, 54 respondents (16.9%) subsequently denied any history of psychological treatment when they completed the demographic items. Nevertheless, participants in both crowdsourcing subsamples reported higher levels of psychopathology than the college students.

We collected a total of 2,107 observations across these five subsamples. We subsequently dropped (a) 146 participants who failed one or more of our four validity checks and (b) 81 participants who failed to complete at least 80% of the items in the main internalizing item pool. This yielded a final total of 1,870 participants in the Development sample (see Table 2 for the final number of observations in each subsample).

Table 2 provides demographic information for the final set of participants in each subsample. Compared to the students, respondents in the two crowdsourcing subsamples were substantially older (for Qualtrics, M age = 42.4 years; for Prolific, M age = 30.8 years) and had a higher proportion of men (44.8% in the Qualtrics subsample, 43.7% in the Prolific subsample).

Validation Sample

Assessment protocol. We will test the generalizability of our findings using responses collected from a separate Validation sample (N = 502) that was recruited via Prolific; these data were collected in July 2020. These participants completed the 198 items contained in the 35 non-sexual dysfunction scales that were created during the Development phase (to be described subsequently); the sexual dysfunction items were not assessed in this sample. In addition, participants completed the Neuroticism (48 items) and Extraversion (41 items) domain scales from the Faceted Inventory of the Five-Factor Model (FI-FFM); Watson, Nus, & Wu, 2019). The FI-FFM items are rated on a 5-point scale ranging from *strongly disagree* to *strongly agree.*

Participants. The Validation sample was recruited as part of a study examining the overlap between internalizing and somatoform psychopathology. It consisted of two subsamples. First, 252 participants endorsed a screening item indicating that they had visited a doctor's office for a possible neurological problem at some point in their lifetime. The remaining 250 participants reported a lifetime history of psychological treatment. An inspection of their Prolific ID numbers revealed that 6 participants also were included in the Development sample; these participants were dropped, yielding a final sample of 496 non-overlapping respondents.

Table 2 reports demographic information for this final set of participants. It is noteworthy that the two Prolific groups were very similar with regard to age (mean age = 30.8 and 31.7 years in the Development and Validation samples, respectively), gender (54.7% and 57.3% women, respectively), and treatment status (current treatment = 26.4% and 30.2%, respectively; past treatment history = 80.5% and 78.8%, respectively).

Scale Development Analyses

Overview

All scale-level analyses used Pearson correlations, whereas all item-level analyses employed polychoric correlations to account for the ordinal nature of the data. We begin this section by discussing the development of preliminary internalizing spectrum scales based on the 395 items included in the main section of the assessment protocol. As discussed earlier, the use of screening questions introduced systematic patterns of missingness into the sexual dysfunction items. Consequently, we describe the development of the sexual dysfunction scales separately.

Main Internalizing Analyses

Preliminary analyses. As noted previously, we dropped participants in the Development sample who failed to complete at least 80% of the items in the main internalizing section. The large majority of the remaining participants had no missing data. For those with some missing data, we used Proc MI in SAS 9.4 (with 11 imputations) to complete the missing responses.

The main scale development analyses were conducted between June and September 2019. Our first major decision concerned how best to use data from the various Development subsamples. We began by aggregating the responses from the two crowdsourcing (N = 646) and three student subsamples (N = 1,224). Next, we created two random student subsamples using Proc Surveyselect in SAS 9.4. At this point, we had three subsamples of roughly equal size. We then conducted a series of analyses to determine whether these three subsamples produced similar findings. These analyses revealed that the subsamples produced results that were broadly similar to one another, while also demonstrating some non-trivial differences.

To illustrate this key point, we report principal factor analyses of the 52 non-sexual HICs in each subsample, using squared multiple correlations as the initial communality estimates. In each case, we extracted four factors and rotated them to oblique simple structure using Promax. The loadings from these solutions are presented in Supplemental Tables S2 through S4. Three

factors—which we labeled Distress (e.g., Depressed Mood, Numbing, Nightmares), Fear (e.g., Situational Phobia, Public Spaces Anxiety), and Mania (e.g., Euphoric Mood, Grandiosity)—appeared to replicate well across all three subsamples, whereas the fourth factor did not.

We quantified factor similarity by computing comparability coefficients (Everett & Entrekin, 1980; Finn, 1986). Comparability coefficients involve deriving regression-based factor scoring weights for each solution (Everett & Entrekin, 1980; Finn, 1986; Gorsuch, 1983), which then can be applied to every dataset. Everett (1983) suggested that comparability coefficients ≥ .90 indicate that the same factors emerged across solutions. In our data, comparability coefficients ranged from .950 to .992 (median = .974) for the Distress factor, from .953 to .989 (median = .972) for the Fear factor, and from .916 to .992 (median = .958) for the Mania factor. Thus, these three factors replicated quite well across all three subsamples. In contrast, comparability coefficients for the fourth factor ranged from only .683 to .984 (median = .808), indicating poor replicability across solutions.

Based on these preliminary analyses, we concluded that the findings from the three subsamples were sufficiently similar to justify combining them into a single overall analysis. We therefore aggregated the responses into a single dataset (N = 1,870) for all subsequent analyses.

Primary analyses. The main scale development process proceeded in four basic steps. The overall pool of 395 items was too large to be subjected to item-level factor analyses. Accordingly, following the general data analytic plan developed by the HiTOP Measurement Workgroup, the first step was to conduct structural analyses based on the rational HICs. The goal here was to identify smaller subsets of closely related constructs that then could be subjected to more manageable item-level analyses. Once again, we conducted a principal factor analysis of the 52 non-sexual HICs, this time using the overall sample. Given that our goal was to divide our large item pool into more manageable subdomains, we sought to identify the maximum number

of interpretable factors. This analysis revealed four meaningful factors.

The Promax loadings from this four-factor solution are presented in Table 3. Based on their core content, these four factors can be labeled Distress, Fear, Social/Somatic Anxiety, and Mania, respectively. Based on their loadings, we assigned the HICs to corresponding subdomains; note that because of the large size of the Distress factor, HICs with salient loadings on other factors (e.g., Well-Being) were assigned to other subdomains. At this point, the Distress subdomain was defined by 26 HICs representing 194 items; Fear included 11 HICs representing 86 items; Social/Somatic Anxiety was marked by 9 HICs representing 68 items; Mania was defined by 6 HICs representing 47 items.

The Distress subdomain still was larger than desirable for item-level analyses. We therefore subjected the 26 Distress HICs to an additional principal factor analysis. This analysis revealed two meaningful subfactors, which we labeled Core Distress and Panic/PTSD. The Promax loadings from this analysis are shown in Table 4. Core Distress consisted of 14 HICs representing 104 items; Panic/PTSD was defined by 12 HICs representing 90 items.

Based on these initial results, we had carved our item pool in five subdomains, ranging in size from 47 items (Mania) to 104 items (Core Distress). The second step in the scale development process was to conduct a series of item-level factor analyses within each of these subdomains. Consistent with the general data analytic plan developed by the HiTOP Measurement Workgroup, we sought to identify the maximum number of factors/scales that could be differentiated in our data. We began with an overall analysis of the items included within each subdomain. This initial analysis was supplemented by structural analyses of smaller subgroups of items to determine whether broader dimensions could be decomposed into more specific factors. These analyses produced a provisional set of 35 scales.

The third step focused on improving the discriminant validity of the scales by eliminating problematic items. This involved (a) identifying highly correlated scales [defined as correlating ≥ .75] and (b) factoring their items in a single joint analysis; items with salient loadings on both factors were candidates for elimination at this stage. At the conclusion of this stage, the 35 provisional scales included a total of 258 items.

It should be noted that discriminant validity was a much greater concern in some areas than in others. This can be illustrated using the HICs that were included in each subdomain. The mean correlations (after *r*-to-*z* transformation) of the HICs within the Mania, Fear and Social/Somatic Anxiety subdomains were .50, .50, and .54, respectively. In contrast, the mean correlations within the two subdomains carved from the initial Distress factor were much higher: .65 for Panic/PTSD and .72 for Core Distress. The Core Distress subdomain clearly represented the most general problem vis-à-vis discriminant validity. In fact, seven of its constituent HICs—Anhedonia, Anxious Mood, Depressed Mood, Lassitude, Numbing, Worry, and Worthlessness/Guilt—had correlations ranging from .71 to .89, with a mean value of .81.

The fourth step was to refine the scales further by eliminating redundant items wherever possible, while keeping a minimum of 3 items per scale. Item redundancy was operationally defined as polychoric correlations ≥ .80. A total of 60 items were dropped due to redundancy, producing a final set of 198 items. After the completion of this final stage, no scale correlations were ≥ .75. Only four item pairs (out of a total of 19,503 item correlations) had polychoric correlations ≥ .80, all of them in the Excoriation and Trichotillomania scales.

Thus, the scale development process ultimately yielded 35 scales consisting of 198 items. The final scales are listed in Table 5 (along with the sexual dysfunction scales, which are discussed in the following section). Supplemental Table S5 displays the item composition of the

scales vis-à-vis the original rational HICs. Most of the scales essentially represent reduced versions of the original HICs (see Table S5). However, nine scales are more complex. Five scales are blends of two HICs: Agoraphobia combines items from Enclosed Spaces Anxiety and Public Spaces Anxiety; Body Dissatisfaction includes content from Body Dissatisfaction and Body Preoccupation; Cognitive Problems contains items from Cognitive Problems and Hyperactive Cognition; Social Anxiety subsumes content from Interactive Anxiety and Evaluative Anxiety; and Traumatic Reactions combines items from Traumatic Avoidance and Traumatic Intrusions. The four remaining scales—Anhedonic Depression, Anxious Worry, Euphoric Energy, and Grandiosity—contain items from three or more HICs.

Sexual Dysfunction Analyses

Analyses of the sexual dysfunction items were conducted in September/October 2019. As noted earlier, these items contained systematic patterns of missingness. This patterned missingness was handled using full information maximum likelihood estimation for item-level analyses, and the mean of non-missing responses for scale-level analyses. In addition to conducting analyses using the overall sample, we also analyzed men and women separately, given sex differences in each of the assessed domains of sexual function (e.g., Basson, 2000).

Preliminary analyses of the five rational HICs indicated a one-factor solution in the full sample (N = 1,862) and for the male (N = 585) and female (N= 1,256) subsamples separately, although low orgasmic function was a weak indicator of the factor ($\lambda s \leq .21$ versus $\lambda s > .6$ for all other indicators), consistent with its weak correlations ($r < .30$) with the other domains. Low orgasmic function was consequently treated as an orphan HIC and analyzed separately.

We conducted item-level analyses of the low sexual desire, difficulties with arousal, sexual pain, and distress related to sexual dysfunction items (N = 27) in an EFA framework,

treating them as categorical indicators using WLSMV estimation and Geomin rotation; parallel analysis (Horn, 1965) was based on treating the items as continuous indicators using MLR estimation. A total of 1,862 participants were included in these analyses, as 8 respondents (0.4%) of the full sample were missing responses on all items.

Parallel analysis indicated four factors in the full sample, which also corresponded to the most interpretable solution. We named these factors Distress and Difficulties with Arousal (e.g., "I felt guilty about sexual difficulties"; "It was hard for me to become sexually aroused"), Low Sexual Desire (e.g., "I easily lost interest in sexual activity"), Solo Sexual Desire (e.g., "I wanted to engage in sexual activity by myself'), and Sexual Pain (e.g., " I experienced unwanted pain during sex"). Highly similar factors emerged in the analyses of both women and men. After removing redundant items (again, defined as polychoric correlations ≥ .80), we created four initial scales: Distress Related to Sexual Dysfunction (4 items), Low Sexual Desire (4 items), Solo Sexual Desire (3 items), and Sexual Pain (4 items). However, we subsequently dropped the Solo Sexual Desire scale, as it was weakly related to the other scales and essentially unrelated to BFI-2-S Negative Emotionality, suggesting that it was capturing neither psychopathology in general, nor difficulties with sexual functioning specifically.

The orgasmic function items (N = 8) were analyzed in the same way, based on the subset of participants who had responded to at least one of them (N = 1,538). Parallel analysis indicated two factors in the full sample, which corresponded with two interpretable factors that converged well in the male (N = 554) and female (N = 966) samples. We labeled these factors Low Orgasmic Function (e.g., "I found it difficult to reach orgasm") and (2) Pleasurable Orgasm (e.g., "My orgasms were pleasurable"). Each factor had three clear markers in all three solutions. However, paralleling the situation with Solo Sexual Desire, subsequent analyses indicated that

Pleasurable Orgasm was non-pathological in nature, given that it was weakly (and sometimes negatively) correlated with the other scales and actually had a significant negative correlation with BFI-2-S Negative Emotionality. It therefore was dropped. Consequently, the scale development process produced four scales—Distress Related to Sexual Dysfunction, Low Sexual Desire, Sexual Pain, and Low Orgasmic Function—with a total of 15 items (see Table 5).

Scale Properties

Internal Consistency

In this section, we report results that help to establish the nature and correlates of these preliminary internalizing scales. We begin by examining their internal consistency. Table 5 presents average interitem correlations (AICs) for each scale in the Development and Validation samples; these values are presented as Pearson correlations, which are most commonly reported in the assessment literature.

Assessment experts recommend that scale items should be moderately correlated with one another. Specifically, Clark and Watson (1995, 2019) recommend AICs in the .15 to .50 range. Clark and Watson (1995, 2019) further emphasize that the optimal AIC depends on the generality versus specificity of the assessed construct: Scales assessing very broad dimensions ideally should have AICs toward the bottom of this range, whereas measures of narrower constructs should have AICs toward the top of this range. Given that we sought to identify the maximum number of dimensions that clearly could be differentiated in our data (i.e., we opted for splitting rather than lumping), most of our scales should tap narrow constructs and have AICs closer to the top of this range.

In the Development sample, 15 of 39 scales (38.5%) had AICs within this optimal range (values ranged from .377 to .492), and an additional eight scales (20.5%) had slightly higher

AICs ranging from .507 to .549. Overall, the AICs ranged from .377 to .687, with a median value of .533. In the Validation sample, 15 of 35 scales (42.9%) had AICs within the optimal range (values ranged from .279 to .487), and additional eight scales (22.9%) had slightly higher AICs ranging from .502 to .549. Overall, the AICs ranged from .279 to .728 in this sample, with a median value of .507. As expected, these results indicate that most of these preliminary scales assess narrow dimensions reflecting specific forms of psychopathology.

Finally, it is worth noting that some scales consistently had much higher AICs than others. We quantified this assertion by computing a column-vector correlation for the 35 scales assessed in both samples. This vector correlation was .71, indicating a strong level of consistency across the two samples.

Internal Structure

General factor solution. Next, we examined the internal structure defined by the scales. These structural analyses were designed to answer two questions. First, do the scales define a single general factor? If so, this would provide some preliminary evidence that they all are indicators of the same overarching domain (viz., internalizing). We examined this issue by conducting a separate principal factor analysis of the scales included in each sample.

Table 6 presents the loadings from these general factor solutions. These results confirm the existence of a strong and broad general factor in both samples. Thirty-three scales (84.6%) and 34 scales (97.1%) had loadings ≥ .50 in the Development and Validation samples, respectively. Indicators of distress and negative emotionality tended to be particularly strong markers of this general dimension. For example, Anhedonic Depression, Shame/Guilt, Anxious Worry, Angry Hostility, and Separation Insecurity all had loadings ranging from .77 to .80 in the Development sample, and from .72 to .79 the Validation sample. Panic, agoraphobia, PTSD, and OCD

symptoms also were strong markers of this factor.

We examined the consistency of this general factor by computing Tucker's (1951) congruence coefficient (*K*), based on the loadings for the 35 scales that were common to both solutions. Lorenzo-Seva and ten Berge (2006) suggest that *K* values ≥ .95 indicate good factor similarity, whereas values between .85 and .94 represent fair similarity. In our data, *K* = .994, establishing that the same general factor emerged in both samples.

Four-factor solution. Second, do these scales define meaningful subdimensions? If so, this would help to explicate the nature of the basic types of content subsumed within the domain. We examined this issue using the 35 scales that were common to both samples. We conducted separate principal factor analyses in each solution. Four interpretable and well-defined factors emerged in each sample and were rotated using Promax.

Table 7 presents the loadings from these four-factor solutions (the factor correlations are reported in Supplemental Table S6). It is noteworthy that these factors broadly resemble those that were obtained earlier from the structural analyses of the rational HICs (see Table 3). The largest factor can be labeled Distress and it clearly resembles the broad first factor shown in Table 3. This factor is defined by scales reflecting subjective distress and negative emotionality—such as Anhedonic Depression, Anxious Worry, Traumatic Reactions, Suicidality, Shame/Guilt, Angry Hostility, and Physiological Panic—that also tended to be the strongest markers of the general factor (see Table 6). We labeled the second factor Body Dysmorphia. It is similar to the third factor in Table 3 and had four consistent markers—Body Dissatisfaction, Body Focus, Perfectionism, and Appetite Gain—with loadings > .40 in both solutions. The third factor represents Fear (similar to the second factor in Table 3); it had six consistent markers—Situational Phobia, Animal Phobia, Cleaning, Blood-Injection Phobia, Agoraphobia, and

Compulsive Rituals—across samples. The fourth factor is Mania, and essentially represents the same dimension as the last factor in Table 3. It had five consistent markers—Euphoric Energy, Grandiosity, Decreased Need for Sleep, Recklessness, and Well-Being—across solutions.[1]

We computed comparability coefficients to quantify the level of factor similarity across samples. The Distress, Fear, and Mania factors had comparability coefficients ranging from .956 to .978 (median = .974), thereby demonstrating an impressive level of structural stability. In contrast, coefficients for the Body Dysmorphia factor were only .865 (Development sample) and .877 (Validation sample), representing a more modest level of replicability. These lower coefficients reflect the fact that this factor was broader in the Development sample (nine scales had loadings ≥ .40) than in the Validation sample (only four scales had loadings ≥ .40).

Correlations with Personality

Finally, we examine relations with neuroticism/negative emotionality (N/NE) and extraversion. At the diagnostic level, previous research has established that internalizing disorders are related to multiple FFM traits, but show the strongest and most consistent links to N/NE (Goldberg, Krueger, Andrews, & Hobbs, 2009; Kotov et al., 2017; Watson et al., in press; Widiger et al., 2019). For example, Kotov, Gamez, Schmidt, and Watson (2010) reported meta-analytic associations between FFM traits and 10 depressive and anxiety disorder diagnoses. These diagnoses were most strongly related to N/NE (mean r = .39), but also showed moderate negative associations with both conscientiousness (mean r = -.28) and extraversion (mean r = -.25); they were largely unrelated to openness (mean r = -.09) and agreeableness (mean r = .01). Based on these results, we would expect indicators of internalizing to be positively

[1] We also conducted an additional principal factor analysis in the Development sample that included the four sexual dysfunction scales. These scales defined a separate fifth factor; the other factors closely resembled those shown in Table 7.

correlated with N/NE and—to a lesser extent—negatively related to extraversion.

At the same time, however, it is clear that some internalizing symptoms are more strongly related to N/NE than others. For instance, Watson and Naragon-Gainey (2014) examined associations between FFM domain scores and both self-reported and interview-based symptoms of anxiety, depression, and mania. They found that N/NE correlated strongly with symptoms of general distress/negative emotionality (e.g., depressed mood, anxious mood, worry); more moderately with symptoms of social phobia, panic, PTSD, and OCD; and only modestly with other types of symptoms, including agoraphobia and specific phobia.

Table 8 presents correlations between the internalizing scales and N/NE and extraversion domain scores based on the BFI-2-S in the Development sample and the FI-FFM in the Validation sample. We will begin by noting that the results were highly consistent across samples and measures. We quantified this point by computing column-vector correlations for the 35 scales common to both samples. These vector correlations were .97 for N/NE and .93 for extraversion, establishing a very strong level of replicability.

Table 8 indicates that the internalizing scales varied widely in their associations with N/NE. Consistent with previous research, distress-based scales that are strong markers of the general Internalizing factor tended to have particularly substantial links to this trait. Seven scales had correlations ≥ .50 with N/NE in both samples, with coefficients ranging from .52 to .75: these included Anxious Worry, Anhedonic Depression, Separation Insecurity, Shame/Guilt, Social Anxiety, Angry Hostility, and Cognitive Problems. In marked contrast, five scales—Decreased Need for Sleep, Trichotillomania, Blood-Injection Phobia, Euphoric Energy, and Grandiosity—consistently were weakly correlated with N/NE, with coefficients ranging from only -.12 to .27 across samples. More generally, indicators of mania, specific phobia, the

OCRDs, and sexual dysfunction tended have relatively weak associations with N/NE.

Most scales had relatively modest associations with extraversion. Consistent with the findings of Watson and Naragon-Gainey (2014), extraversion was most strongly negatively related to Social Anxiety, with correlations of -.62 and -.45 in the Development and Validation samples, respectively. It also was moderately negatively correlated with Anhedonic Depression (*r*s = -.48 and -.25, respectively). Furthermore, as would be expected based on previous research (Watson, Stanton, Khoo, Ellickson-Larew, & Stasik-O'Brien, 2019), some scales consistently displayed positive associations with extraversion. These included Well-Being (*r*s = .48 and .59, respectively) and some indicators of mania (for Grandiosity, *r*s = .18 and .44, respectively; for Euphoric Energy, *r*s = .15 and .34, respectively).

Future Directions

We have described the development of 39 preliminary internalizing scales, which were created during Phase 1 of the HiTOP measure development project. As we discussed at the beginning of this paper, the development of these preliminary scales is part of a larger effort by the HiTOP consortium to create a comprehensive self-report measure of psychopathology. In Phase 2, our 213 retained items will be combined with those from the other spectra (viz., somatoform, thought disorder, detachment, and disinhibited and antagonistic externalizing) to further the development of this comprehensive measure.

Many of our preliminary scales will undergo significant modification in response to these Phase 2 analyses, and some may disappear altogether. As one example, the Somatoform Group—which, as discussed previously, was tasked with the assessment of eating pathology—has created a preliminary Body Image and Weight Concerns scale. This scale contains content (e.g., "I was dissatisfied with the shape of my body," "I thought other people's bodies looked

better than mine") that appears quite similar to that contained in our Body Dissatisfaction scale (e.g., "I was dissatisfied with my weight," "I spent a lot of time comparing my appearance to that of others"). It seems very likely that these items will be collapsed into a single scale in Phase 2. As another example, the Detachment and Thought Disorder Groups both created preliminary Anhedonia scales that contain content that clearly overlaps with many of the items in our Anhedonic Depression scale. All of these items may be collapsed into a single scale as a result of the Phase 2 analyses.

The Phase 2 analyses also will help to clarify where our items/scales ultimately fall within the hierarchical structure of psychopathology. To be sure, many of our preliminary scales clearly represent core manifestations of internalizing. Scales such as Anxious Worry, Anhedonic Depression, Separation Insecurity, Shame/Guilt, and Cognitive Problems are (a) strongly correlated with one another, (b) substantial markers of the general Internalizing factor (see Table 6), and (c) strongly correlated with N/NE (see Table 8). Generally speaking, these types of scales—which also are strong indicators of the Distress subdimension (see Table 7)—are likely to define internalizing in subsequent phases of the HiTOP assessment project.

In contrast, other scales performed more equivocally, such that their ultimate placement remains uncertain. These scales include indicators of mania, specific phobia, sexual dysfunction, and the OCRDs. It will be interesting to see where these scales/items are located within a comprehensive symptom-based structure of psychopathology.

References

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

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

Basson, R. (2000). The female sexual response: A different model. *Journal of Sex &Marital Therapy, 26,* 51-65.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7,* 309-319. doi: 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,* 1412-1427. doi: 10.1037/pas0000626

Cutshall, C., & Watson, D. (2004). The Phobic Stimuli Response Scales: A new self-report measure of fear. *Behaviour Research and Therapy, 42,* 1193-1201. doi 10.1016/j.brat.2003.08.003

Dornbach-Bender, A., Ruggero, C. J., Waszczuk, M. A., Gamez, W., Watson, D., & Kotov, R. (2017). Mapping emotional disorders at the finest level: Convergent validity and joint structure based on alternative measures. *Comprehensive Psychiatry*, *79,* 31–39. doi:10.1016/j.comppsych.2017.06.011

Everett, J. E. (1983). Factor comparability as a means of determining the number of factors and their rotation. *Multivariate Behavioral Research, 18,* 197–218. doi: 10.1207/s15327906mbr1802_5

Everett, J. E., & Entrekin, L. V. (1980). Factor comparability and the advantages of multiple group factor analysis. *Multivariate Behavioral Research, 2,* 165–180. doi: 10.1207/s15327906mbr1502_4

Finn, S. E. (1986). Structural stability of the MMPI in adult males. *Journal of Consulting and Clinical Psychology, 54,* 703–707. doi: 10.1037/0022-006X.54.5.703

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. . (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment,* 14, 485-496. doi: 10.1037.1040-3590.14.4.485

Forbes, M. K., Baillie, A. J., & Schniering, C. A. (2016a). A structural equation modeling analysis of the relationships between depression, anxiety, and sexual problems ove time. *Journal of Sex Research, 53,* 942-954. doi: 10.1080.00224499.2015.1063576

Forbes, M. K., Baillie, A. J., & Schniering, C. A. (2016b). Should sexual problems be included in the internalising spectrum? A comparison of dimensional and categorical models. *Journal of Sex and Marital Therapy, 42,* 70–90. doi:10.1080/0092623X.2014.996928

Forbush, K. T., South, S. C., Krueger, R. F., Iacono, W. G., Clark, L. A., Keel, P. K, Legrand, L. N., & Watson, D. (2010). Locating eating pathology within an empirical diagnostic taxonomy: Evidence from a community-based sample. *Journal of Abnormal Psychology, 119,* 282-292. doi: 10.1037/a0019189

Forbush, K. T., & Watson, D. (2013). The structure of common and uncommon mental disorders. *Psychological Medicine, 43,* 97-108. doi: 10.1017/S0033291712001092

Goldberg, D. P., Krueger, R. F., Andrews, G., & Hobbs, M. J. (2009). Emotional disorders: Cluster 4 of the proposed meta-structure for *DSM-V* and *ICD-11. Psychological Medicine, 39,* 2043-2059. doi: S0033291709990298

Gootzeit, J., Markon, K., & Watson, D. (2015). Measuring dimensions of PTSD: The Iowa Traumatic Response Inventory. *Assessment. 22,* 152-166. doi: 10.1177/1073191114540747

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hogan, R. T. (1983). A socioanalytic theory of personality. In M. Page (Ed.), *1982 Nebraska Symposium on Motivation* (pp. 55-89). Lincoln: University of Nebraska Press.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-185. doi: 10.1007/BF02289447

Johnson, R. E., Rosen, C. C., & Djurdjevic, E. (2011). Assessing the impact of common method variance on higher order multidimensional constructs. *Journal of Applied Psychology, 96,* 744-761. doi: 10.1037/a0021504

Koffel, E. (2011). Further validation of the Iowa Sleep Disturbances Inventory. *Psychological Assessment, 23,* 587-598. doi: 10.1037/a0022818

Koffel, E., Khawaja, I. S., & Germain, A. (2016). Sleep disturbances in posttraumatic stress disorder: Updated review and implications for treatment. *Psychiatric Annals, 46,* 173-176. doi: 10.3928/00485713-20160125-01

Koffel, E., Krebs, E. E., Arbisi, P. A., Erbes, C. R., & Polusny, M. A. (2015). The unhappy triad: Pain, sleep complaints, and internalizing symptoms. *Clinical Psychological Science, 4,* 96-106. doi: 10.1177/2167702615579342

Koffel, E., & Watson, D. (2009). The two-factor structure of sleep complaints and its relation to depression and anxiety. *Journal of Abnormal Psychology, 118,* 183-194. doi: 10.1037/a0013945

Kotov, R., Gamez, W., Schmidt, F. L., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin, 136,* 768-821. doi: 10.1037/a0020327

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A

dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126,* 454-477. doi:10.1037/abn0000258

Levin, R., & Nielsen, T. A. (2007). Disturbed dreaming, posttraumatic stress disorder, and affect distress: A review and neurocognitive model. *Psychological Bulletin, 133,* 482-528. doi: 10.1037/0033-2909.133.3.482

Lorenzo-Seva, U., & ten-Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, 2,* 57-64. doi: 10.1027/1614-2241.2.2.57

Mahoney, A. E. J., Hobbs, M. J., Newby, J. M., Williams, A. D., Sunderland, M. (2016), & Andrews, G. The Worry Behaviors Inventory: Assessing the behavioral avoidance associated with generalized anxiety disorder. *Journal of Affective Disorders, 203,* 256-264. doi: 10.1016/j.jad.2016.06.020

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17,* 437-455. doi: 10.1037/a0028085

Reio, T. G., Jr. (2010). The threat of common method variance bias to theory building. *Human Resource Development Review, 9,* 405-411. doi:10.1177/1534484310380331

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E.,…Zimmermann, J. (2019). Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. *Journal of Consulting and Clinical Psychology, 87,* 1069–1084. doi: 10.1037/ccp0000452

Ruggero, C. J., Kotov, K., Watson, D., Kilmer, J. N., Perlman, G., & Liu, K. (2014). Beyond a single index of mania symptoms: Structure and validity of subdimensions. *Journal of Affective Disorders, 161,* 8-15. doi: 10.1016/j.jad.2014.4.02.044

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory-2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality, 68,* 69-81. doi: 10.1016/j.jrp.2017.02.004

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies.* (Personnel Section Report No. 984). Washington, DC: Department of the Army.

Waszczuk, M. A., Kotov, R., Ruggero, C., Gamez, W., & Watson, D. (2017). Hierarchical structure of emotional disorders: From individual symptoms to the spectrum. *Journal of Abnormal Psychology, 126,* 613–634. doi:10.1037/abn0000264

Watson, D. (2012). Objective tests as instruments of psychological theory and research. In H. Cooper (Ed.), *Handbook of research methods in psychology. Vol. 1: Foundations, planning, measures, and psychometrics* (pp. 349-369). Washington, DC: American Psychological Association. doi: 10.1037/13619-019

Watson, D., Levin-Aspenson, H. F., Waszczuk, M. A., Conway, C. C., Dalgleish, T., Dretsch, M. N.,…HiTOP Utility Workgroup. (in press). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry.*

Watson, D., & Naragon-Gainey, K. (2014). Personality, emotions, and the emotional disorders. *Clinical Psychological Science, 2,* 422-442. doi: 10.1177/2167702614536162

Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R.,…Ruggero, C. J. (2012). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment, 19,* 399-420. doi: 10.1177/1073191112449857

Watson, D., Stanton, K., Khoo, S., Ellickson-Larew, S., & Stasik-O'Brien, S. M. (2019). Extraversion and psychopathology: A multilevel hierarchical review. *Journal of Research in Personality, 81,* 1-10. doi:10.1016/j.jrp.2019.04.009

Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015). Explicating the psychopathological correlates of anomalous sleep experiences. *Psychology of Consciousness: Theory, Research, and Practice, 2,* 57-78. doi: 10.1037/cns0000038

Watson, D., Stasik-O'Brien, S. M., Ellickson-Larew, S., & Stanton, K. (2018). Explicating the dispositional basis of the OCRDs: A hierarchical perspective. *Journal of Psychopathology and Behavioral Assessment, 40,* 497-513.doi: 10.1007/s10862-018-9667-5

Watson, D., & Wu, K. D. (2005). Development and Validation of the Schedule of Compulsions, Obsessions, and Pathological Impulses (SCOPI). *Assessment, 12,* 50-65. doi: 10.1177/1073191104271483

Widiger, T. A., Sellbom, M., Chmielewski, M., Clark, L. A., DeYoung, C. G., Kotov, R.,…Wright, A. G. C. (2019). Personality in a hierarchical model of psychopathology. *Clinical Psychological Science, 7,* 77-92.doi: 10.1177/2167702618797105

Yufik, T., & Simms, L. J. (2010). A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms. *Journal of Abnormal Psychology, 119,* 764-776. doi: 10.1037/a0020981

Table 1

The 57 Original Target Constructs in the Initial Item Pool

Target Construct	Sample Item
Anxiety (12 constructs, 93 items)	
Animal phobia (7)	I was afraid of spiders
Anxious mood (6)	I felt anxious
Blood-injection phobia (8)	I felt faint at the sight of blood
Chronic, excessive worry (6)	I worried about almost everything
Enclosed spaces anxiety (8)	I avoided small, tight, or enclosed spaces
Interactive anxiety (9)	I was uncomfortable meeting new people
Performance anxiety (8)	I got nervous because others were evaluating me
Physiological panic (9)	My heart was racing or pounding
Psychological panic (8)	I got so worked up my mind went blank
Public spaces anxiety (8)	I was afraid of getting trapped in a crowd
Situational phobia (9)	I was afraid of heights
Worry behaviors (7)	I made lists to reduce my anxiety
Sleep (3 constructs, 22 items)	
Insomnia (6)	I had trouble falling asleep
Lassitude (9)	I had trouble getting out of bed
Nightmares (7)	I was troubled by nightmares
Depression (10 constructs, 77 items)	
Anhedonia (9)	Nothing seemed interesting to me
Appetite gain (6)	I felt like eating more than usual
Appetite loss (6)	I had a sudden loss in appetite
Cognitive problems (7)	I had trouble concentrating
Depressed mood (7)	I felt down and discouraged
Psychomotor agitation (6)	I had trouble sitting still
Psychomotor retardation (6)	I talked more slowly than usual
Suicidality (11)	I thought about killing myself
Well-Being (9)	I felt that I had a lot to look forward to
Worthlessness/guilt (10)	I felt like a failure
PTSD (5 constructs, 34 items)	
Avoidance (6)	I avoided people who might bring back bad memories
Hypervigilance (7)	I felt “on guard” and on edge
Intrusions/re-experiencing (8)	I had repeated memories of a traumatic event
Numbing (7)	I felt distant or cut off from other people
Peritraumatic dissociation (6)	The world seemed strange and “different”

(table continues)

Table 1 (cont.)

Target Construct	Sample Item
OCD (5 constructs, 40 items)	
Checking/doubting (8)	I spent a lot of time checking and rechecking things
Counting/numbers (6)	I avoided numbers that could bring bad luck
Obsessing (8)	I repeatedly had immoral thoughts
Ordering/rituals/symmetry (9)	I felt compelled to follow certain rituals
Washing/cleaning (9)	I was obsessed with cleanliness
OCRDs (5 constructs, 38 items)	
Body dissatisfaction (7)	I wished I could change certain parts of my body
Body preoccupation (9)	I spent a lot of time looking at myself
Excoriation (7)	I felt a strong urge to pick my skin
Hoarding (8)	I collected things that I did not need
Trichotillomania (7)	I coped with stress by pulling my hair
Mania (7 constructs, 52 items)	
Decreased need for sleep (6)	I felt like I could go for days without sleeping
Emotional lability (7)	I had sudden, intense mood swings
Euphoric mood/excessive energy (7)	I felt very excited and hyper for no reason
Grandiosity/overconfidence (7)	I thought I was on a mission from a higher power
Hyperactive cognition (7)	I spoke so fast that people could not follow me
Increased goal-directed activity (7)	I set many unrealistic goals for myself
Recklessness (11)	I spent money recklessly
Maladaptive Traits (5 constructs, 39 items)	
Irritability/hostility (10)	Everyone got on my nerves
Nonsuicidal self-injury (8)	I cut myself on purpose
Perfectionism (6)	I demanded perfectionism from myself
Perseveration (7)	I got fixated on doing things a certain way
Separation insecurity (8)	I worried that others would abandon me
Sexual dysfunction (5 constructs, 35 items)	
Difficulties with arousal (6)	I felt uninterested during sexual activity
Low orgasmic function (8)	I was unable to reach orgasm
Low sexual desire (8)	I had little or no interest in sexual activity
Sexual pain (6)	I experienced unwanted pain during sex
Sex-related distress (7)	I felt sexually inadequate

Note. PTSD = posttraumatic stress disorder. OCD = obsessive-compulsive disorder. OCRDs = obsessive-compulsive and related disorders.

Table 2

Demographic Breakdown of the Development and Validation Samples

	Development Sample					
Variable	Qualtrics	Prolific	UNT	UND	UTS	VALID
Number of Participants	328	318	580	510	134	496
Mean Age	42.4	30.8	20.7	19.4	19.0	31.7
% Women	54.6	54.7	77.9	71.4	69.4	57.3
% Men	44.8	43.7	21.0	27.8	27.6	38.7
% Other or missing	0.6	1.6	1.0	0.8	3.0	4.0
% White	77.7	84.6	58.8	75.7	11.2	74.4
% Black/African-American	10.7	2.8	12.9	4.1	6.0	5.0
% Asian	4.0	3.8	6.9	8.8	61.9	8.7
% Other or missing	7.6	8.8	21.4	11.4	21.0	11.9
% Hispanic or Latino/a	10.7	5.7	28.1	15.1	--*	13.5
% Non-Hispanic or Latino/a	88.7	94.0	71.0	84.5	--	85.9
% Missing	0.6	0.3	0.9	0.4	--	0.6
% Currently in treatment	100.0	26.4	16.2	10.8	8.2	30.2
% Not currently in treatment	0.0	70.8	83.1	88.4	88.9	69.8
% Missing	0.0	2.8	0.7	0.8	3.0	0.0
% With past treatment history	94.5	80.5	43.3	28.4	20.1	78.8
% With no past treatment	5.5	18.9	56.2	70.1	77.6	21.2
% Missing	0.0	0.6	0.5	1.4	2.2	0.0

Note. UNT = University of North Texas. UND = University of Notre Dame. UTS = University of Toronto-Scarborough. VALID = Validation Sample.

*This question was not asked in this sample.

Table 3

Promax Factor Loadings of the 52 Rational HICs

HIC	I	II	III	IV
Depressed Mood[a]	**.91***	-.15	.20	-.16
Suicidality[a]	**.90***	.03	-.17	.00
Anhedonia[a]	**.89***	.02	.04	-.11
Numbing[a]	**.88***	-.05	.14	-.07
Worthlessness[a]	**.79***	-.11	.30	-.09
Peritraumatic Dissociation[a]	**.78***	.06	.03	.10
Obsessing[a]	**.76***	.07	.04	.15
Emotional Lability[a]	**.75***	-.10	.17	.16
Traumatic Intrusions[a]	**.73***	.14	-.01	.08
Lassitude[a]	**.71***	-.07	.29	-.11
Psychological Panic[a]	**.67***	.25	.05	.04
Worry[a]	**.66***	-.04	.39	-.08
Hostility[a]	**.66***	.07	.09	.13
Psychomotor Retardation[a]	**.64***	.24	.01	.09
Anxious Mood[a]	**.64***	-.03	**.42**	-.11
Insomnia[a]	**.63***	.08	.07	-.01
Nonsuicidal Self-Injury[a]	**.62***	.25	-.30	.22
Hyperactive Cognition[a]	**.60***	-.09	.32	.28
Cognitive Problems[a]	**.59***	-.06	**.41**	.04
Psychomotor Agitation[a]	**.57***	.04	.26	.17
Traumatic Avoidance[a]	**.55***	.17	.13	.08
Nightmares[a]	**.55***	.29	-.11	.07
Physiological Panic[a]	**.54***	.32	.10	.06
Appetite Loss[a]	**.53***	.06	-.05	.17
Hypervigilance[a]	.36*	.34	.28	.05
Perseveration[a]	.34*	.19	.28	.22
Situational Phobia[b]	-.03	**.77***	.09	.00
Enclosed Spaces Anxiety[b]	.24	**.69***	-.03	-.05
Animal Phobia[b]	-.17	**.69***	.10	.07
Blood-Injection Phobia[b]	-.00	**.67***	.06	.06
Cleaning[b]	-.03	**.66***	.08	.15
Counting[b]	.04	**.56***	.03	.29
Public Spaces Anxiety[b]	.37	**.56***	.16	-.27
Trichotillomania[b]	.13	**.48***	-.16	.26
Ordering[b]	.07	**.47***	.19	.29
Hoarding[b]	.29	.35*	.11	.23
Excoriation[b]	.15	.30*	.02	.17

(table continues)

Table 3 (cont.)

HIC	I	II	III	IV
Body Preoccupation[c]	-.03	-.02	**.69***	.23
Body Dissatisfaction[c]	.21	.05	**.61***	-.03
Worry Behaviors[c]	.25	.07	**.58***	.15
Perfectionism[c]	.06	.00	**.50***	.36
Separation Insecurity[c]	**.49***	-.05	**.44**	.09
Interpersonal Anxiety[c]	.32	.33	**.42***	-.29
Checking[c]	.21	.26	**.42***	.15
Evaluative Anxiety[c]	.25	**.42***	**.41***	-.19
Appetite Gain[c]	.10	.13	.38*	.08
Grandiosity/Overconfidence[d]	-.15	.16	.02	**.77***
Euphoric mood/Excessive Energy[d]	.01	.10	.05	**.76***
Increased Goal-Directed Activity[d]	-.14	-.04	**.41**	**.68***
Well-Being[d]	**-.74***	-.02	.19	**.62**
Recklessness[d]	.38	.09	-.02	**.58***
Decreased Need for Sleep[d]	.33	.16	-.08	**.55***

Note. N = 1,870. Loadings ≥ .40 are in bold. * = Highest loading (absolute value within ± .01) in row.

[a] Included in the Distress subdomain

[b] Included in the Fear subdomain

[c] Included in the Social/Somatic Anxiety subdomain

[d] Included in the Mania Subdomain

Table 4

Promax Loadings of the 26 Distress HICs

Scale	I	II
Depressed Mood[a]	**.89***	.04
Anxious Mood[a]	**.88***	.02
Lassitude[a]	**.85***	.00
Worry[a]	**.85***	.06
Cognitive Problems[a]	**.83***	.05
Worthlessness[a]	**.79***	.14
Numbing[a]	**.74***	.23
Anhedonia[a]	**.69***	.24
Hyperactive Cognition[a]	**.65***	.24
Emotional Lability[a]	**.60***	.29
Psychomotor Agitation[a]	**.57***	.31
Hostility[a]	**.45***	**.41**
Perseveration[a]	**.43***	.35
Insomnia[a]	**.42***	.35
Nonsuicidal Self-Injury[b]	-.12	**.82***
Nightmares[b]	.01	**.76***
Traumatic Intrusions[b]	.18	**.74***
Obsessing[b]	.33	**.62***
Psychological Panic[b]	.35	**.61***
Suicidality[b]	.26	**.60***
Traumatic Avoidance[b]	.26	**.59***
Physiological Panic[b]	.36	**.55***
Psychomotor Retardation[b]	.38	**.52***
Peritraumatic Dissociation[b]	**.42***	**.51***
Hypervigilance[b]	.39	**.48***
Appetite Loss[b]	.27	.37*

Note. N = 1,870. Loadings ≥ .40 are in bold. * = Higher loading (absolute value within ± .01) in row.

[a] Included in the Core Distress subdomain

[b] Included in the Panic/PTSD subdomain

Table 5

Average Interitem Correlations of the Preliminary Internalizing Spectrum Scales

Scale	Development	Validation
Agoraphobia (6)	.598	.505
Angry Hostility (9)	.533	.512
Anhedonic Depression (10)	.542	.449
Animal Phobia (5)	.393	.442
Anxious Worry (7)	.574	.549
Appetite Gain (4)	.594	.642
Appetite Loss (3)	.549	.507
Blood-Injection Phobia (3)	.483	.604
Body Dissatisfaction (5)	.611	.635
Body Focus (6)	.440	.429
Checking (7)	.552	.597
Cleaning (8)	.465	.392
Cognitive Problems (4)	.591	.580
Compulsive Rituals (6)	.475	.487
Decreased Need for Sleep (5)	.556	.545
Distress Related to Sexual Function (4)	.569	.—
Euphoric Energy (7)	.484	.550
Excoriation (3)	.670	.483
Grandiosity (7)	.418	.458
Hoarding (6)	.431	.382
Indecisiveness (3)	.687	.728
Insomnia (4)	.591	.522
Low Orgasmic Function (3)	.434	.—
Low Sexual Desire (4)	.538	.—
Nightmares (3)	.642	.606
Nonsuicidal Self-Injury (3)	.484	.515
Perfectionism (5)	.492	.507
Perseveration (5)	.476	.279
Physiological Panic (6)	.507	.415
Recklessness (7)	.377	.443
Separation Insecurity (8)	.420	.454
Sexual Pain (4)	.608	.—
Shame/Guilt (4)	.652	.647
Situational Phobia (5)	.383	.352
Social Anxiety (10)	.598	.554
Suicidality (3)	.579	.536
Traumatic Reactions (8)	.533	.482
Trichotillomania (3)	.527	.483
Well-Being (10)	.526	.502

Note. N = 1,870 (Development sample), 496 (Validation Sample). Values shown represent Pearson correlations. The number of items for each scale is in parentheses.

Table 6

Loadings of the Internalizing Scales on a General Factor

Scale	Development	Validation
Physiological Panic	**.81**	**.75**
Traumatic Reactions	**.80**	**.77**
Anhedonic Depression	**.80**	**.72**
Shame/Guilt	**.79**	**.76**
Checking	**.77**	**.79**
Angry Hostility	**.77**	**.77**
Anxious Worry	**.77**	**.74**
Separation Insecurity	**.77**	**.79**
Hoarding	**.75**	**.74**
Perseveration	**.75**	**.70**
Cognitive Problems	**.74**	**.68**
Agoraphobia	**.73**	**.70**
Social Anxiety	**.72**	**.64**
Suicidality	**.69**	**.72**
Compulsive Rituals	**.69**	**.74**
Nightmares	**.67**	**.63**
Insomnia	**.67**	**.56**
Cleaning	**.66**	**.62**
Indecisiveness	**.66**	**.60**
Situational Phobia	**.65**	**.66**
Recklessness	**.64**	**.73**
Decreased Need for Sleep	**.64**	**.59**
Nonsuicidal Self-Injury	**.63**	**.63**
Appetite Loss	**.57**	**.61**
Body Dissatisfaction	**.56**	**.59**
Euphoric Energy	**.54**	**.67**
Body Focus	**.54**	**.66**
Perfectionism	**.54**	**.61**
Trichotillomania	**.53**	**.56**
Animal Phobia	**.52**	**.60**
Blood-Injection Phobia	**.50**	**.57**
Distress Related to Sexual Function	**.50**	.—
Appetite Gain	**.50**	**.61**
Excoriation	**.47**	**.55**
Grandiosity	**.45**	**.60**
Low Sexual Desire	**.44**	.—
Sexual Pain	**.42**	.—
Low Orgasmic Function	.31	.—
Well-Being	-.20	.01

Note. N = 1,475 (Development sample), 496 (Validation sample). Loadings with an absolute value ≥ .40 are in bold.

Table 7

Promax Loadings of the Internalizing Scales in the Four-Factor Solution

	Distress		Body Dysmorphia		Fear		Mania	
Internalizing Scale	DEV	VAL	DEV	VAL	DEV	VAL	DEV	VAL
Anhedonic Depression	**.88**	**.92**	.13	-.08	-.04	-.10	-.07	.13
Suicidality	**.75**	**.49**	-.11	-.15	.05	.11	.13	**.41**
Insomnia	**.67**	**.52**	.10	.06	-.04	.02	.07	.10
Traumatic Reactions	**.65**	**.51**	.13	.06	.06	.16	.13	.23
Anxious Worry	**.63**	**.79**	**.41**	.12	-.07	.07	-.07	-.06
Shame/Guilt	**.61**	**.80**	.38	.20	-.08	-.07	.04	.04
Angry Hostility	**.59**	**.45**	.13	-.00	.09	.15	.14	.35
Cognitive Problems	**.55**	**.76**	**.42**	.08	-.11	-.11	.04	.11
Physiological Panic	**.53**	**.45**	.17	-.08	.23	.27	.06	.28
Appetite Loss	**.52**	.30	-.07	-.08	-.01	-.03	.29	**.54**
Nightmares	**.50**	.29	-.02	.03	.22	.35	.13	.12
Nonsuicidal Self-Injury	**.48**	.11	-.22	-.06	.18	.01	.37	**.71**
Social Anxiety	**.45**	**.67**	.36	.10	.25	.18	-.26	-.15
Perseveration	.32	.21	.30	.27	.20	.15	.14	.28
Well-Being	**-.81**	**-.71**	.20	.31	.00	.01	**.52**	**.46**
Body Dissatisfaction	.08	**.45**	**.76**	**.54**	.01	.07	-.15	-.26
Body Focus	-.22	.08	**.71**	**.58**	.02	-.01	.28	.25
Indecisiveness	.35	**.71**	**.52**	.16	-.04	-.04	-.03	-.07
Separation Insecurity	**.43**	**.64**	**.51**	.23	-.03	-.00	.05	.13
Perfectionism	.01	.22	**.50**	**.50**	-.01	-.07	.25	.17
Appetite Gain	.01	.25	**.47**	**.44**	.18	.03	-.03	.10
Checking	.27	.34	**.41**	.34	.22	.30	.08	.06
Situational Phobia	.10	.02	.02	-.07	**.72**	**.79**	-.03	.07
Animal Phobia	-.12	-.08	.10	.11	**.67**	**.74**	-.02	-.01
Cleaning	.04	-.07	.06	.23	**.64**	**.61**	.12	.03
Blood-Injection Phobia	-.05	-.05	-.02	-.09	**.61**	**.54**	.09	.29
Agoraphobia	**.40**	.19	-.02	-.09	**.56**	**.69**	-.07	.07
Compulsive Rituals	.08	.02	.07	.14	**.52**	**.40**	.25	.38
Trichotillomania	.14	.04	-.09	-.00	.39	.20	.27	**.45**
Hoarding	.29	.20	.14	.24	.32	.07	.22	**.43**
Excoriation	.08	.13	.10	.11	.29	.01	.16	**.45**
Euphoric Energy	-.05	-.05	.08	.11	.07	-.01	**.73**	**.79**
Grandiosity	-.18	-.24	.01	.12	.20	.11	**.70**	**.78**
Decreased Need for Sleep	.32	-.04	-.06	-.03	.02	-.00	**.63**	**.80**
Recklessness	.30	.13	.06	.03	-.02	.04	**.58**	**.71**

Note. N = 1,870. Loadings ≥ |.40| are in bold. DEV = Development sample. VAL = Validation sample.

Table 8

Correlations of the Internalizing Scales with Neuroticism/Negative Emotionality and Extraversion

	Neuroticism/Negative Emotionality		Extraversion	
Internalizing Scale	DEVEL	VALID	DEVEL	VALID
Anxious Worry	**.68***	**.75***	-.32	-.19
Anhedonic Depression	**.63***	**.67***	**-.48***	-.25
Separation Insecurity	**.59***	**.70***	-.29	-.06
Shame/Guilt	**.56***	**.68***	-.34	-.17
Social Anxiety	**.55***	**.56***	**-.62***	**-.45***
Angry Hostility	**.52***	**.61***	-.23	.06
Cognitive Problems	**.52***	**.61***	-.32	-.14
Traumatic Reactions	**.49***	**.56***	-.28	-.03
Indecisiveness	**.47***	**.60***	-.31	-.24
Suicidality	**.46***	**.51***	-.33	-.04
Body Dissatisfaction	**.46***	**.54***	-.23	-.17
Physiological Panic	**.46***	**.53***	-.30	-.04
Insomnia	**.45***	**.50***	-.31	-.07
Agoraphobia	**.40***	**.42***	-.30	-.03
Checking	**.40***	**.55***	-.27	.04
Nightmares	.38	**.43***	-.20	.04
Perseveration	.38	**.41***	-.24	.11
Distress Related to Sexual Function	.36	.—	-.25	.—
Hoarding	.33	**.42***	-.24	.12
Appetite Loss	.29	.32	-.19	.08
Appetite Gain	.27	**.40***	-.16	.00
Situational Phobia	.27	.36	-.23	.01
Nonsuicidal Self-Injury	.26	.24	-.17	.20
Low Sexual Desire	.26	.—	-.25	.—
Low Orgasmic Function	.23	.—	-.12	.—
Perfectionism	.22	.37	.02	.12
Compulsive Rituals	.20	.35	-.13	.13
Sexual Pain	.20	.—	-.14	.—
Recklessness	.19	.30	-.02	.31
Body Focus	.18	.38	.03	.22
Cleaning	.18	.33	-.17	.10
Excoriation	.17	.27	-.13	.12
Animal Phobia	.16	.31	-.14	.11
Decreased Need for Sleep	.15	.19	-.04	.27
Trichotillomania	.14	.23	-.12	.09
Blood-Injection Phobia	.12	.23	-.09	.14
Euphoric Energy	.02	.27	.15	.34
Grandiosity	-.12	.10	.18	**.44***
Well-Being	**-.51***	-.37	**.48***	**.59***

Note. N = 1,534 to 1,868 (development sample), 496 (validation sample). Correlations with an absolute value ≥ .40 are highlighted. DEVEL = Development sample. VALID = Validation sample.