Integrating HiTOP and Computational Psychiatry for a New Era of Clinical Science

Timothy A. Allen[1], Alison M. Schreiber[2], Scott D. Blain[3], Christina Carlisi[4], Colin G. DeYoung[5], Alex Fornito[6], Claire M. Gillan[7], Vina M. Goghari[8], Nathan T. Hall[9], Philipp Homan[10], Antonia N. Kaczkurkin[11], Roman Kotov[12], Robert F. Krueger[5], Robert D. Latzman[13], Aysenur Okan[9], Jennifer L. Tackett[14], Michael N. Hallquist[9+], & the HiTOP Neurobiological Foundations Workgroup

[+]Senior Author

[1] University of Pittsburgh, Pittsburgh, PA, USA
[2] University of Kentucky, Lexington, KY, USA
[3] The Ohio State University, Columbus, OH, USA
[4] University College London, London, England
[5] University of Minnesota-Twin Cities, Minneapolis, MN, USA
[6] Monash University, Clayton, Victoria, Australia
[7] Trinity College Dublin, Dublin, Ireland
[8] University of Toronto, Toronto, ON, Canada
[9] University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[10] University of Zurich, Zurich Switzerland
[11] Vanderbilt University, Nashville, TN, USA
[12] Stony Brook University, Stony Brook, NY, USA
[13] Otsuka Pharmaceutical Development & Commercialization, Inc., Princeton, NJ, USA
[14] Northwestern University, Evanston, IL, USA

Corresponding Author: Timothy A. Allen, 100 North Bellefield Ave., Pittsburgh, PA, 15237; allenta@pitt.edu

Abstract

For decades, advancements in understanding and treating mental illness have been hindered by a categorical psychiatric nosology that fails to describe and explain how symptoms emerge and relate to one another. In response, two separate initiatives have contributed to a renewed sense of optimism for scientific and translational discovery. First, quantitative models of psychopathology improve clinical description by relying on empirical data to provide a more accurate representation of the structure of mental illness. Chief among these approaches is the Hierarchical Taxonomy of Psychopathology (HiTOP), which organizes psychopathology based on patterns of symptom covariation observed across numerous studies. In parallel, computational psychiatry seeks to identify neurocognitive processes that give rise to psychopathology and leverage them to forecast important clinical and functional outcomes. Here, we highlight points of convergence and complementarity between HiTOP and computational psychiatry and propose further integration. HiTOP provides an empirically based approach to understanding the structure of psychopathology, but lacks strong explanations of how symptom dimensions are connected to underlying neurocognitive processes. Conversely, computational psychiatry's emphasis on cognitive processes and multivariate prediction make it well-suited to linking lower level dynamics with symptom variability. Yet, many contemporary computational psychiatry findings lack specificity and validity as a result of a continued reliance on categorical diagnoses. We conclude by highlighting potential barriers that will need to be addressed to maximize the productivity of future collaborative efforts between these initiatives.

For much of its history, psychiatry and clinical psychology have struggled with the fundamental question of how best to classify mental illness. Early descriptive approaches gave way to the categorical diagnostic system articulated in the *Diagnostic and Statistical Manual - 5th edition* (*DSM-5*), but this system has faced criticism for failing to capture how symptoms are organized, poor correspondence with biological systems, and limited clinical validity (Cuthbert & Insel, 2013; Kotov et al., 2017; Ruggero et al., 2019; Stephan & Mathys, 2014). In response, two major research movements – quantitative psychopathology and computational psychiatry – have sought to address these challenges from different angles.

Quantitative models of psychopathology use statistical methods to derive higher-order dimensions based on covariation between narrower symptoms and traits. The most comprehensive of these empirically based models is the Hierarchical Taxonomy of Psychopathology (HiTOP), which provides a consensus view of the structure of psychopathology. HiTOP's quantitative approach to organizing psychopathology has yielded symptom dimensions that are more reliable and valid than classic diagnoses (Conway et al., 2019; Kotov et al., 2020, 2021; Krueger et al., 2021; Ruggero et al., 2019; Watson et al., 2022). However, strong theories that explain how HiTOP dimensions connect to biological and cognitive processes are currently lacking.

Computational approaches adopt a different strategy, using formal, mathematical models to identify the neurocognitive processes disrupted in mental illness and data-driven algorithms to predict important outcomes from clinical and computational measures. Some early computational psychiatrists (Montague et al., 2012), and biological psychiatrists before them, hoped such advances could transform our

approach to clinical *description*, shifting the focus to objective indicators rather than signs and symptoms. Yet, the search for “biomarkers” or “computational phenotypes” that can be used to diagnose psychopathology has so far yielded little success, in part because our understanding of mental illness is defined primarily by self-reports of emotional and behavioral distress, as well as societal conceptions of psychological well-being (Maddux et al., 2019). As a result, many modern computational studies have shifted their focus away from revising psychiatric taxonomy and instead concentrated on predicting or explaining existing categorical diagnostic constructs, despite their known shortcomings.

Thus far, HiTOP and computational psychiatry have evolved as separate research frameworks, each responding to distinct flaws in the current nosology. Here, we consider the benefits of integrating the two approaches. HiTOP offers computational psychiatry an opportunity to target more precise and coherent descriptions of clinical phenomena. Conversely, computational psychiatry can provide HiTOP with the methodological tools needed to link underlying neurocognitive dynamics to the structure of psychopathology and its outcomes.

HiTOP: An Alternative Nosology

HiTOP is based on a quantitative approach to nosology, in which symptoms and traits that frequently co-occur are organized into common higher-order dimensions; those that co-occur less often are organized into separate dimensions. For example, many mood and anxiety symptoms co-occur within individuals and can be understood as reflecting a broader dimension of internalizing psychopathology (Watson et al., 2022). Historically, quantitative approaches have contributed to a number of useful

assessment measures (Achenbach, 1966; Kay et al., 1987), as well as broader conceptual models of personality, cognition, affect, and psychopathology (Clark & Watson, 1991; Costa & McCrae, 1992; Krueger et al., 2012; McGrew, 2009). HiTOP extends these prior efforts in three critical ways.

First, HiTOP aims to be comprehensive, striving to represent all signs and symptoms of psychopathology (though, see Michelini et al., 2024), whereas prior efforts have been constrained to specific symptom domains or forms of psychopathology. The model includes a general factor of psychopathology (known as a superspectrum) at the highest level, as well as narrower spectra (internalizing, disinhibited externalizing, antagonistic externalizing, thought disorder, detachment, and provisionally, somatoform), subfactors, syndromes, and eventually individual traits and symptoms (Figure 1a). Higher-order dimensions, particularly the six spectra, capture major features of psychopathology, such as internalizing, that are not specifically described by categorical systems or are fragmented across diagnoses (Forbes, Neo, et al., 2024). Lower-order subfactors, symptoms, and traits provide detailed coverage of the more narrow features of psychopathology (e.g., fear and distress subfactors of internalizing), enabling more fine-grained phenotypic description.

Second, the model synthesizes a large body of research: 261 studies informed its original structure, with additional studies providing evidence for its utility and validity (Kotov et al., 2022). More recently, a meta-analysis of factor analytic research using *DSM* diagnoses, which included over 120,000 participants and 35 studies, found that a model closely resembling HiTOP was a strong fit to the data, providing a quantitative validation of the framework (Ringwald et al., 2023). Another bottom-up effort to model

patterns of covariation among *DSM-5* symptoms similarly yielded a structure that closely resembled the HiTOP spectra, suggesting that the spectra are implicitly “baked in” to *DSM* categories (Forbes, Baillie, et al., 2024).

Third, whereas prior quantitative efforts focused on measure development or conceptual advancements (goals which HiTOP also pursues), HiTOP explicitly aims to become a formal nosology, including efforts to optimize the framework for clinical settings (Kotov et al., 2022; Ruggero et al., 2019). To support this aim, the consortium has published tutorials and workshops on clinical integration (Hopwood et al., 2018; Su et al., 2025), developed a standardized measure to guide assessment (Simms, 2023), created cross-walks to translate HiTOP profiles into billing codes (Ruggero et al., 2019) and shown that clinicians generally view HiTOP more favorably than the *DSM* (Balling et al., 2023).

The Two Cultures of Computational Psychiatry

Computational psychiatry is an interdisciplinary field that uses analytic tools and conceptual frameworks from cognitive neuroscience, computer science, and decision science to understand the mental computations that contribute to psychopathology, aiming to uncover brain-behavior relationships that traditional methods miss (Hallquist et al., 2024). The field encompasses two research approaches, or “cultures,” that vary primarily by their objective (though see Kind & Dayan, 2024 for a further subdivision): a data-driven approach uses machine learning methods to optimize the prediction of an outcome or set of outcomes from a set of input features, whereas a theory-driven approach uses generative models, instantiated in formal equations, to mathematically explain relations between inputs to the model (e.g., sensory or perceptual experience)

and outputs (e.g., behavior or neural activity; Bennett et al., 2019; Huys, Maia, et al., 2016).

Data-Driven Computational Psychiatry

Data-driven computational psychiatry uses machine learning methods to leverage the power of big data to parse heterogeneity in psychopathology, explain its underlying mechanisms, and forecast its outcomes[1]. Unlike traditional hypothesis-driven research, which links a few variables to an outcome, machine learning leverages many input variables (i.e., features), often drawn from diverse data sources, to maximize the predictive power of an algorithm (Chekroud et al., 2021) (Figure 1b).

Broadly, machine learning methods can be categorized into two categories, unsupervised and supervised, with hybrid methods also possible. Unsupervised approaches aim to uncover latent structures in data without relying on labeled outcomes or explicit supervision, allowing the model to identify previously unknown patterns, groupings, or relationships. These models represent an important point of convergence between HiTOP and data-driven computational psychiatry because factor analysis – the primary tool used in quantitative studies of psychopathology – is an unsupervised procedure for identifying latent factors that explain patterns of covariation (e.g., among symptoms). Other examples of unsupervised methods include clustering algorithms, principal component analysis, and mixture modeling. By reducing complex data into fewer dimensions or subtypes, these methods reveal underlying structures that can inform classification or prediction, sometimes in combination with supervised learning techniques.

[1] We use the term forecast to refer to the estimation of future values, such as treatment outcomes. In contrast, we use prediction to refer to the estimation of any outcome based on observed data, such as estimating a dimension score based on a set of biological features.

In contrast to unsupervised approaches, supervised learning methods involve an algorithm learning how features of a training dataset relate to an outcome variable provided by a knowledgeable “supervisor” (Sutton & Barto, 1998). For instance, researchers might train algorithms to learn patterns between a set of input features (e.g., whole-brain beta coefficients from fMRI) and expert judgments of diagnosis or treatment outcome. Critically, whereas traditional methods focus primarily on linear effects, machine learning algorithms are capable of incorporating nonlinear and interactive effects among variables in high-dimensional spaces, even when the interpretability of those effects is not straightforward. As a result, the process that *generates* the data is not always easily inferred in data-driven approaches, which can limit their explanatory value (though increasingly, the field is developing ways around this limitation; J. Chen et al., 2023).

Recent innovations in machine learning have also led to the emergence of large language models (LLMs) capable of processing and generating human-like text. These models are trained using a self-supervised learning paradigm, in which a basic unit of text (i.e., a word or token) is hidden, and the model learns to predict it based on surrounding context. This approach may enable clinically meaningful insights about disease presentation and progression from unstructured text, including electronic health records and clinician notes. Though still emerging, LLMs offer a powerful new tool for leveraging big data, particularly when integrated with more precise and reliable descriptive frameworks like HiTOP.

Theory-Driven Computational Psychiatry

While machine learning approaches prioritize predictive accuracy, theory-driven applications of computational psychiatry place a stronger emphasis on understanding the data-generating process that produces behavior (Figure 1c). Many different types of models can fall under the umbrella of “theory-driven” computational psychiatry, including reinforcement learning models, sequential sampling models (e.g., the popular drift diffusion model), ideal Bayesian observer models, and dynamic causal modeling (DCM), though this list is by no means exhaustive. Critically, all of these models are “generative” in the sense that they aim to formally specify the latent (or hidden) states and dynamics that map inputs to relevant outputs, including neural activity or observed behavior (Hallquist et al., 2024; Stephan et al., 2015). Researchers can test alternative hypotheses by adjusting components or assumptions of the model and observing the effects on simulated behavior. This process of *formalizing* competing models distinguishes theory-driven computational modeling from traditional psychological theories, which often rely on loosely specified verbal hypotheses that are difficult to operationalize or falsify (Smaldino, 2020).

To illustrate, consider that many psychological theories link anhedonia and depression to impaired reward sensitivity (Pizzagalli, 2014; Treadway & Zald, 2013) – a poorly defined construct that could encompass distinct processes, including deficits in reward learning, valuation, or choice behavior. Reinforcement learning (RL) models provide a framework to disambiguate these possibilities by describing how agents learn which actions yield desirable outcomes through experience. RL models typically consist of three components: a prediction error, a learning rule, and a choice rule.

The prediction error drives learning, and is defined as the difference between the outcome the agent receives and the value that was expected:

$$\delta_t = r_t - Q_t(a_t)$$

Here, r_t is the reward on a trial *t*, $Q_t(a_t)$ is the expected value of the chosen action a_t, and δ_t signals whether the outcome was better or worse than expected. Prediction errors can update expectations about how valuable an action is according to a simple learning rule:

$$Q_{t+1}(a_t) = Q_t(a_t) + \alpha\,\delta_t$$

where the learning rate parameter, α, controls how much new outcomes shift expectations. A high learning rate implies rapid adjustment, whereas low values imply conservative updating. In anhedonia, one hypothesis is that outcomes carry less weight, resulting in reduced learning. In RL models, this would be captured by a negative association between anhedonia and the learning rate (Chase et al., 2010; Dombrovski et al., 2010).

Alternative hypotheses can also be considered within this framework. For example, learning deficits may be reward-specific, which can be modeled with separate learning rates for gains and losses (Brown et al., 2021; Mukherjee et al., 2020), or they may reflect valuation differences, such that the impact of rewards is subjectively diminished (Huys et al., 2013), even if learning is intact. The latter possibility can be captured by introducing a reward scaling parameter, ρ, into the equation above, with lower values indicating reduced subjective reward.

Finally, RL models must translate values into actions via a choice rule, which often uses a softmax function:

$$P(a) = \frac{exp(\beta Q_t(a))}{\Sigma_{a'} exp(\beta Q_t(a))}$$

The softmax rule converts estimated values ($Q_t(a)$) into probabilities, ensuring that all possible actions have some chance of being selected. Actions with higher expected values are assigned higher probabilities, making them more likely, but not guaranteed. The inverse temperature parameter, β, controls value sensitivity: high β produces consistent selection of the best option, while low β produces more exploratory or variable choices. Some studies have found that anhedonia is associated with lower β, suggesting that the core difficulty may lie not in learning per se, but in translating learned values into decisions (Pike & Robinson, 2022).

Once hypotheses are formalized within a set of models, researchers can fit them to observed data by estimating parameters that minimize the discrepancy between predicted and actual behavior – a process known as model inversion. Models can then be compared based on their ability to predict behavioral patterns (e.g., learning task reversals) or other levels of data (e.g., BOLD activity). This approach allows researchers to infer the latent processes and parameters underlying behavior (Stephan et al., 2015). Notably, empirical findings for anhedonia have thus far been inconsistent, and one reason for this may be that studies have relied on case-control designs (e.g., depression vs. healthy controls) to study what is inherently a dimensional phenotype. Moving forward, the field would benefit from larger samples, more elegant models, and most importantly, a stronger descriptive framework, to better characterize the computational mechanisms of anhedonia (Hallquist et al., 2024).

What HiTOP Offers Computational Psychiatry

Improved Phenotypic Precision

In the past 30 years, personal navigation has improved dramatically as a function of modern software applications that integrate location information with mapping data denoting the presence of roads and bridges or even real-time traffic conditions. Critically, the accuracy of these applications ultimately depends on the quality of their inputs. Even with precise location information, inaccurate or incomplete road data can lead to inefficient routes or unexpected roadblocks. This is not unlike the current state of affairs in contemporary psychiatry and clinical psychology, where an invalid nosology – an inaccurate view of the terrain of psychopathology – can limit the impact of even our most powerful computational techniques.

Accumulating evidence points to important inaccuracies in the *DSM*-based nosology. Meta-analyses examining hundreds of taxometric studies have shown that symptoms vary in a continuous fashion rather than exhibiting the “zones of rarity” (Kendell & Jablensky, 2003) that would be expected if diagnostic thresholds were scientifically valid (Haslam et al., 2012; 2020). Boundaries between healthy and pathological functioning are often arbitrary: key symptoms of major disorders are present in many individuals who lack a diagnosis, and the relationship between symptoms and outcomes typically follows a continuous gradient (Ruscio, 2019). Boundaries between disorders are similarly blurry, with patients frequently meeting criteria for multiple disorders at rates far exceeding chance, suggesting most psychiatric disorders are not discrete entities (Caspi et al., 2020; Kessler et al., 2005; Krueger & Markon, 2006).

Overlooking fundamental flaws in the current diagnostic system hampers the impact of modern computational research. For example, many computational studies

compare model-based parameters or high-dimensional feature sets between individuals with and without a given diagnosis. In doing so, they implicitly assume that diagnostic labels reflect discrete and valid categories that can be distinguished according to some underlying neural or computational process. Consequently, models often fail not because of the quality of the computational approach, but because the diagnostic labels they target do a poor job of carving nature at its joints (Frenay & Verleysen, 2014; Scott & Nelson, 2024). Indeed, recent evidence suggests classification errors in predictive models can often be attributed to overlapping symptoms and high rates of comorbidity (Olshan et al., 2024). Poor reliability (i.e., high measurement error) of diagnoses (Clark et al., 2017; Regier et al., 2013) also weakens associations with behavioral or neural measures and increases variability in observed effect sizes, undermining replicability (Gell et al., 2024; Nikolaidis et al., 2022). This may be one reason that meta-analyses of learning and decision-making differences across clinical groups often report vast heterogeneity in findings (Halahakoon et al., 2020; Lloyd et al., 2024; Pike & Robinson, 2022).

HiTOP addresses these problems by offering computational psychiatrists a more valid, reliable framework for characterizing individual differences. Its variable-centered approach describes individuals along a set of continuously varying dimensions (DeYoung et al., 2022), allowing psychopathology to be assessed using continuous measures that have greater statistical power and reliability relative to categorical diagnoses (Carcone et al., 2015; MacCallum et al., 2002; Markon et al., 2011; see Simms, 2023 for details on the development of a HiTOP self-report measure). It also encourages dimensional sampling, in which participants are selected across the full

range of symptom severity rather than being dichotomized into groups. This retains valuable information from individuals with subthreshold symptoms, minimizes the risk of model misspecification, and improves generalizability (Conway & Brown, 2018; Kotov et al., 2018; Preacher et al., 2005). Collectively, these improvements can lead to larger effect sizes when linking psychopathology to external variables (DeYoung, Hilger, et al., 2024). Indeed, one recent review found that HiTOP dimensions explained nearly three times more variance than traditional diagnoses in concurrent and prospective validators, including clinical status, functioning, service utilization, and biomarkers (Kotov et al., 2022).

By enforcing artificial group boundaries, the *DSM*-based nosology not only undermines the reliability and validity of findings but also renders them less interpretable. In any computational study in which a group difference emerges on some model-based parameter, that difference may plausibly reflect: 1) overall severity, such that any diagnosis would show the same difference on the parameter when compared to controls (cf. differential deficit concept of Chapman & Chapman, 1978); 2) a specific feature of the diagnosis, as opposed to the entire diagnosis; 3) a specific symptom dimension that co-occurs in other disorders, such that many diagnoses, if sampled, would show the same pattern. Group-based designs are incapable of disentangling these possibilities, which may explain why the same learning or cognitive-based deficits are often observed across seemingly distinct conditions (Wise et al., 2023). In reality, these deficits are better attributed to a single symptom or trait dimension that *spans* disorders (Garbusow et al., 2022; Gillan et al., 2016; Smith et al., 2021; Weigard & Sripada, 2021).

A unique strength of HiTOP's focus on dimensions is that it leverages shared variance among individual signs and symptoms to reveal higher-order constructs of import, reframing comorbidity as an informative feature rather than a flaw. This allows researchers to more precisely identify the features of psychopathology that drive observed effects, ultimately helping to reveal common mechanisms underlying related symptoms (Kotov et al., 2017). Gillan and colleagues (2016) recently demonstrated the value of a dimensional and hierarchical approach in computational psychiatry by observing that deficits in goal-directed control – our ability to make deliberate choices – occur across disorders involving compulsivity. Using factor analysis of nine questionnaire measures collected in a heterogeneous sample, they derived a higher-order compulsivity factor that predicted impaired goal-directed control in a reinforcement learning task (Figure 2). This association was unique to compulsivity, suggesting that previous results were likely a product of compulsivity's ubiquity across disorders. Moreover, the finding is stronger for compulsivity than it is for the OCD diagnosis itself, consistent with the idea that more reliable dimensional phenotypes yield larger observed effect sizes than diagnoses (Gillan et al., 2020). Subsequent studies have shown that this same transdiagnostic compulsivity dimension is related to weakened neural representations of action–outcome relationships, suggesting the effect may be a function of an impaired internal model of the environment (Dück et al., 2024; Seow et al., 2021). Taken together, these findings illustrate how adopting a dimensional nosology can enhance phenotypic measurement, leading to more robust and precise associations with the underlying neural or cognitive mechanisms that could one day inform interventions.

Construct Validation and Theory-Building

Concerns about case-control designs undermining the precision of findings, or even obscuring effects entirely, have already been at the forefront of recent computational literature. One recent approach, “computational factor modeling” (CFM), addresses these issues by using computational models to test candidate mechanisms underlying transdiagnostic dimensions identified through exploratory factor analysis in large, unselected samples (Wise et al., 2023). Notably, the procedure for deriving these symptom dimensions is ideally[2] no different in a CFM study than in a HiTOP study; the two initiatives advocate for identical sampling and measurement strategies. The key difference is that CFM emphasizes the value of connecting dimensions to underlying neurocognitive processes.

Why, then, should computational psychiatry turn to HiTOP to characterize psychopathology? One reason is that HiTOP provides a broader, empirically-derived structure that allows researchers to make predictions about which dimensions are related to underlying processes *and which are not*, thus supporting the development of stronger psychological theories. Computational psychiatry has long been concerned with the proliferation of weak theories that offer imprecise predictions about poorly defined constructs (Grahek et al., 2021; Guest & Martin, 2021; Kent et al., 2023; Millner et al., 2020). Indeed, a major strength of computational modeling is its ability to formalize such theories by precisely defining core constructs and their relations to one another in concrete mathematical terms, yielding clear, falsifiable hypotheses

[2] In practice, HiTOP studies have often relied on suboptimal diagnostic indicators to derive dimensions of interest when symptom-level data is sparse, whereas CFM, by definition, relies exclusively on symptom-level indicators. Regardless, both initiatives are clear that symptom-level data is superior to diagnostic-level data.

(Smaldino, 2017, 2020). While many modeling studies apply this level of rigor to focal neural and cognitive processes, far fewer extend the same precision to defining and formalizing their phenotypic targets.

HiTOP's hierarchical structure addresses this gap by promoting construct validation – the process of determining whether a measure accurately represents the hypothetical construct it is intended to assess (Clark & Watson, 2019; Cronbach & Meehl, 1955). Construct validation is a critical tool for developing strong psychological theories because it forces researchers to clearly define what their construct is, what it is not, and how it relates to other constructs (Grahek et al., 2021). Within the HiTOP model, constructs are operationalized as predicted patterns of symptom covariation, expressed in individual studies through factor loadings and cross-loadings. These patterns, represented in the form of higher- and lower-order dimensions in the model, serve as explicit hypotheses about how symptoms should relate to one another and to external variables, and at what magnitude. Researchers can test these hypotheses by collecting symptom-level data, examining observed covariance patterns, and comparing them against the theoretical predictions of the model.

Structural relations formalized within the HiTOP model can be leveraged to strengthen the construct validity of computational models, improving the precision and interpretability of findings. Consider the work of Gillan and colleagues (2016) highlighted in the last section. In the HiTOP model, compulsivity and intrusive thoughts load onto the fear subfactor of the internalizing spectrum, alongside symptoms like physiological anxiety, panic, and phobias (Figure 2). Because these additional symptoms were not assessed in the original study, it remains an open question whether the deficit in

goal-directed control is specific to compulsivity and intrusive thoughts, or whether it applies to the broader class of symptoms associated with the fear subfactor (however, see Gillan et al., 2021). Such questions of construct coverage, and discriminant and convergent validity, become more apparent when working from a model of psychopathology that clearly and comprehensively defines relations between both narrow and broad symptom dimensions.

In a construct validation framework, hierarchical relations between symptom dimensions are the core phenomena that computational models must explain. Thus, the computational psychiatrist's task is not merely to link a candidate neural or cognitive process to a given dimension in the model, but to do so precisely, demonstrating that the process explains why the components of that dimension covary (e.g., why internalizing subfactors correlate) and why the dimension remains distinct from other constructs at the same level (e.g., why internalizing is distinct from externalizing; Allen et al., 2020, 2022; DeYoung & Krueger, 2020). HiTOP's hierarchical structure is ideally suited to answering these types of questions, as it allows researchers to employ a "lumping" and "splitting" approach when testing hypotheses (DeYoung, Blain, et al., 2024). For example, HiTOP *lumps* problems related to disinhibition (i.e., impulsivity) and antagonism into a higher-order externalizing dimension, recognizing that these types of behaviors tend to co-occur and may reflect shared mechanisms, but it also *splits* them at a lower-level, acknowledging that they are also influenced by partially distinct mechanisms.

In some instances, interrogating structural relations within the HiTOP model can reveal links between phenotypes and cognitive or biological processes that might be

otherwise overlooked. For example, Wise and Dolan (2020) recently used a reinforcement learning model to examine whether biases in learning from safety and threat were associated with individual differences in anxiety. In initial analyses, the authors found that anxiety was associated with better safety learning and a computationally derived estimate of the likelihood of safety. In subsequent factor analyses of their anxiety measure, they identified two lower-order factors: a somatic anxiety factor similar to the HiTOP fear subfactor, and a cognitive anxiety factor resembling the HiTOP distress subfactor. While somatic anxiety showed a similar pattern to the broader anxiety dimension – it was associated with better safety learning and greater estimates of the likelihood of safety – cognitive anxiety was linked to enhanced threat learning and stronger estimates of uncertainty. Interestingly, the same compulsivity and intrusive thoughts factor identified in the work of Gillan and colleagues (2016) was also related to better safety learning and stronger perceptions of safety, consistent with the idea that compulsivity and somatic anxiety are both part of a higher-order fear dimension. Thus, in this case, a structural distinction in the HiTOP model appears to map directly to distinct cognitive processes involved in aversive learning. Altogether, by leveraging structural features of the HiTOP model, computational research can uncover nuanced effects that might be otherwise missed using descriptive approaches that have more limited coverage of psychopathology or fail to account for its multidimensional and hierarchical structure.

Phenotypic Alignment with Neurobiology

One of the most common criticisms of the DSM, particularly amongst computational researchers, is that diagnoses map poorly to underlying mechanisms

(Montague et al., 2012). Indeed, studies have consistently shown that genetic liability to psychopathology is continuously distributed throughout the population, with the same genetic factors influencing both mild and severe phenotypes, and substantial overlap across traditional diagnostic categories (J. Martin et al., 2018; Smoller et al., 2019). Some evidence suggests HiTOP dimensions may offer better correspondence with underlying biological processes (Latzman & DeYoung, 2020). For instance, many genetic variants are pleiotropic, conferring risk transdiagnostically in a manner that is consistent with higher-order dimensions (Lee et al., 2021). Likewise, both behavioral and molecular genetic studies converge on a hierarchical genetic structure of psychopathology, with broad factors resembling the HiTOP superspectrum, spectra, and subfactors (Kotov et al., 2020; Krueger et al., 2021; Waldman et al., 2020; Waszczuk et al., 2020; Watson et al., 2022; Williams et al., 2024).

Neuroimaging studies complement this genetic evidence, consistently showing that structural and functional brain changes cut across traditional diagnostic boundaries, with fewer disorder-specific effects (Brandl et al., 2022; Elliott et al., 2018; Goodkind et al., 2015; McTeague et al., 2017, 2020; Repple et al., 2023; Romer et al., 2021; Sharma et al., 2017; Sheffield et al., 2017; Taylor et al., 2023). Moreover, the identification of a robust neuroanatomical phenotype for specific diagnostic categories has proven elusive, despite decades of research (Cao et al., 2025). Several higher-order HiTOP dimensions have also been linked to person-specific deviations in brain structure and connectivity (relative to a population-based reference), consistent with biological processes varying in a continuous fashion (Kist et al., 2025; Marquand et al., 2019; Parkes et al., 2021; Segal et al., 2023).

Other evidence suggests that higher-order dimensions are more tractable neuroscientific targets than diagnoses. In head-to-head comparisons between the two, three studies have found that HiTOP dimensions are better predictors of neural variables than diagnoses when both are assessed in the same sample (Kircanski et al., 2018; E. A. Martin et al., 2021; Reininghaus et al., 2019). Likewise, a recent meta-analysis found that disruptions in circuits connecting the amygdala and anterior cingulate cortex are present across internalizing forms of psychopathology (Marusak et al., 2016), and electrophysiological studies have shown that variability in performance monitoring event-related potentials is more closely associated with higher-order dimensions of psychopathology than any individual disorder (Lutz et al., 2021; Pasion et al., 2023; Pasion & Barbosa, 2019; Perkins et al., 2024; Trayvick et al., 2025).

Finally, evidence from comparative neuroscience provides additional support for the correspondence between dimensions and underlying biological processes. Animal models, for instance, typically represent individual symptoms or features of psychopathology, and tend to show limited coherence with formal diagnostic criteria (Hyman, 2007; Nestler & Hyman, 2010). Dimensions of personality and psychopathology have clear roots in neural systems that have been conserved across species (Latzman et al., 2018; Latzman, Freeman, et al., 2015; Latzman, Hecht, et al., 2015), suggesting they may serve as a useful translational bridge between animal and human research. Indeed, basic animal neuroscience has been instrumental in generating insights into the neural bases of several HiTOP spectra, subfactors, and symptoms. One example of this is the early work of Depue and colleagues (Depue et al., 1994; Depue & Collins, 1999; Depue & Lenzenweger, 2015), which drew on rodent

and nonhuman primate research showing that mesolimbic dopamine regulates approach behavior and incentive motivation. This work has informed a model of trait extraversion that has shaped our understanding of several dimensional phenotypes, including detachment and anhedonia. Similarly, animal studies of subcortical defensive circuitry have helped characterize critical amygdala-prefrontal circuits involved in threat reactivity and negative affect, offering new insights into the biological basis of internalizing psychopathology and its subfactors, fear and distress (Gray & McNaughton, 2000; Panksepp, 2004; Panksepp et al., 2011; Shackman et al., 2016).

Why HiTOP Needs Computational Psychiatry

HiTOP is an atheoretical descriptive nosology (though importantly, it is not free of theoretical assumptions; DeYoung et al., 2022). Consequently, most prior research on the HiTOP model has consisted of factor-analytic investigations aimed at clarifying its structural properties, including the number and nature of dimensions across different levels of resolution (though, see Kotov et al., 2022 for a review of existing validity evidence). However, for HiTOP to serve as a viable alternative nosology to *DSM* that can inform research and guide clinical practice*,* it will need to demonstrate that structural distinctions within the model can help us gain traction on the underlying sources of psychopathology (e.g., biology, genetics) and improve our forecasting of clinical outcomes (Blashfield & Draguns, 1976; Forbes, Ringwald, et al., 2024). Computational psychiatry offers a powerful toolkit for accomplishing these goals (see also Allen et al., 2020; Grahek et al., 2021). Specifically, computational approaches can help to (i) identify and formalize the neurocognitive processes that underlie dimensions

of psychopathology; (ii) leverage biological and cognitive data to validate and constrain the structure of the HiTOP model; and (iii) enhance the clinical utility of the model.

Relating Neurocognitive Processes to Dimensional Psychopathology

Regardless of the strength of our descriptive system, progress in preventing and treating psychiatric illness still fundamentally depends on our ability to identify the underlying mechanisms and processes that explain individual differences in symptoms. But how should we study such processes? Historically, cognitive scientists have tried to do so by studying behavior in experimental paradigms, linking performance metrics (e.g., accuracy, false alarms, reaction time) to individual differences in symptoms. Yet, these metrics can be difficult to interpret, often providing only an indirect window into underlying cognitive processes (Miyake et al., 2000). Similarly, neuroscientific studies have linked individual differences in symptoms and traits directly to underlying neural correlates, but rarely specify the cognitive operations those neural correlates support. Absent a process-level account of how neural activity produces psychopathology, we are left with a proliferation of nonspecific findings, with many traits and symptoms linked to the same neural correlates, and disparate neural correlates linked to the same trait or symptom (Allen et al., 2020; Yarkoni, 2015). This approach of connecting features of psychopathology directly to their neural instantiation in the brain is akin to trying to explain how an orchestra produces music solely by analyzing the physical properties of each instrument (Krakauer et al., 2017; Marr, 1982).

Theory-driven models offer a productive way forward by allowing HiTOP researchers to directly investigate the cognitive mechanisms that give rise to higher-order dimensions of psychopathology. For example, executive function deficits

are well-documented across psychiatric disorders, but traditional task-based measures often show limited coherence and poor diagnostic specificity (Weigard & Sripada, 2021). Weigard and colleagues have addressed this issue by using a popular sequential sampling model, known as the drift diffusion model, to show that broad executive dysfunction reflects a core deficit in a model-based computational process known as efficiency of evidence accumulation (EEA), which indexes how effectively individuals gather goal-relevant information to make adaptive decisions in noisy environments (Weigard & Sripada, 2021). Recent work on EEA shows that it represents a formal, model-based measure of executive functioning that demonstrates strong cross-task reliability, reasonable stability over time, and more cohesiveness than traditional performance metrics (Tomlinson et al., 2025; Weigard, Clark, et al., 2021). Moreover, EEA is significantly reduced across psychiatric disorders including ADHD, bipolar disorder, and schizophrenia (Sripada & Weigard, 2021), represented in the functional dynamics of underlying neural networks (Weigard et al., 2024), closely related to dimensional assessments of externalizing psychopathology (Hall et al., 2021; Tomlinson et al., 2025) and associated with prospective outcomes (Weigard, Brislin, et al., 2021). Taken together, this work provides a compelling example of how theory-driven computational psychiatry can help HiTOP begin to understand the latent cognitive processes and neural dynamics that underlie individual differences in broad, higher-order dimensions of psychopathology.

One benefit of focusing on generative processes is that it can help HiTOP researchers to build theories of psychopathology that bridge levels of analysis, connecting phenotype to cognitive process, and cognitive process to neural

instantiation. This yields converging lines of evidence that can increase confidence in one's theory (Love, 2015; Marr, 1982). To illustrate, consider recent work using computational models of social learning to study how humans decide to cooperate with one another – a topic that is central to understanding many interpersonal forms of psychopathology (Allen et al., 2024; Vanyukov et al., 2019). A traditional reinforcement learning perspective on this problem assumes that individuals should cooperate with others only when doing so has led to reward in the past. Yet, many social interactions defy this principle: if you decide not to buy your neighbor a holiday gift and later find they also did not buy you one, you might feel a sense of relief – an unexpectedly positive outcome – despite receiving no reward. Conversely, if they *did* buy you a gift, you may nonetheless have an aversive experience, perhaps being embarrassed about having nothing to give in return. These reactions depend not on reward, but on the accuracy of social predictions (Figure 3a).

Vanyukov and colleagues (2019) formalized this intuition by proposing that, during social interactions, humans' track the value of their interpersonal policy, or their approach toward a social partner, rather than the reward value of their actions. Indeed, in model comparisons, a policy-learning model outperformed a traditional RL model in predicting cooperative decision-making during a multi-round two-player trust game. Further, activity in the striatum increased when participants correctly predicted their partner's behavior – whether cooperative or not – indicating that the brain tracks the accuracy of social predictions rather than simply favoring positive outcomes (Figure 3b). Thus, neural activity provided convergent evidence for the authors' process-level model of cooperative decision-making.

Characterizing the normative processes that underlie cooperative decision-making provides a foundation for studying how their disruption can lead to psychopathology (Cicchetti, 1993; Sroufe, 1990). Building on these earlier findings, Allen and colleagues (2024) showed that callousness and exploitativeness – two facets of the antagonistic externalizing spectrum – differ in how they relate to interpersonal policy learning (Figure 3c). Specifically, in a sample enriched with personality pathology, callousness was related to a weaker encoding of policy learning signals in the brain's default network, whereas the opposite was true for exploitativeness. This suggests that callous individuals struggle, while exploitative individuals excel, at using past experience to optimize their approach during social interactions. Moreover, the findings complement prior research showing that callousness is related to deficits in the ability to reason about the minds and behaviors of others (i.e., mentalization or social cognition), whereas exploitativeness shows the opposite pattern (Allen et al., 2017; Blain et al., 2025). Interestingly, when the authors conducted a more traditional case-control analysis of the same data, participants diagnosed with borderline personality disorder (BPD) showed weaker policy learning signals, mirroring only the callousness findings (despite BPD participants being significantly higher than controls on both traits). This suggests that the diagnostic heterogeneity of BPD may have suppressed the positive association between learning and exploitativeness, demonstrating how phenotypic imprecision can lead to faulty inference. Taken together, this work shows how theory-driven computational models can contribute to more precise, mechanistically-informed models of psychopathology by identifying and characterizing

the cognitive processes that mediate between neural systems and behavioral expressions of psychopathology.

Using Neurocomputational Measures to Validate and Constrain the HiTOP Model

A key limitation of HiTOP and other contemporary nosologies, including *DSM-5*, is that they rely solely on self-report or interview-based indicators of clinical symptoms (Wittchen & Beesdo-Baum, 2018), which can be vulnerable to systematic biases (Pavlova & Uher, 2020) and present a descriptive characterization that is overly reliant on natural language. Thus, a critical question for future HiTOP research is whether support for the model can be found at distinct levels of analysis, and whether lower levels of analysis can one day be used to inform or revise the model itself.

Recent attempts to leverage computational methods to test or validate structural distinctions within the HiTOP model have already yielded some promising results. For example, data-driven studies have consistently shown that poorer structural and functional network connectivity is common across psychiatric disorders, providing evidence in support of a general factor of psychopathology (Elliott et al., 2018; Kebets et al., 2019; Repple et al., 2023). Other evidence corroborates HiTOP spectra and subfactors. Xia and colleagues (2018) recently used a data-driven approach to identify four brain-based dimensions of psychopathology – distress, fear, thought disorder, and externalizing – that closely mirrored HiTOP domains and replicated across discovery and hold-out samples. Similarly, Royer and colleagues (2024) identified three latent components capturing covariance between parent-reports of symptoms and multimodal imaging data, including a general factor of psychopathology that explained the most covariance with imaging features, and more narrow internalizing/externalizing and

neurodevelopmental components, each of which was associated with unique neural signatures. Using kernel ridge regression models to predict internalizing and externalizing dimensions from resting-state functional connectivity data in a large, diverse sample of children, Qu and colleagues (2025) found that connectivity patterns strongly linked to internalizing and externalizing dimensions were more similar *within* those dimensions than between them (see also J. Chen et al., 2022).

Other studies have examined whether structural distinctions found in self-report data can similarly be detected at the computational or cognitive level of analysis. Wise and Dayan (2020) recently applied partial least squares regression to self-reported trait anxiety and computational indices derived from their reinforcement-learning model of threat and safety learning (described above). They extracted two components that were broadly consistent with the lower-order structure of the internalizing spectrum: the first loaded on symptoms consistent with HiTOP's fear subfactor and correlated with indices of safety learning, whereas the second loaded on symptoms aligned with HiTOP's distress subfactor and was related to indices of threat learning. Similarly, Fox et al. (2025) identified dimensions resembling HiTOP fear and distress in factor analyses of self-report data from over 7,500 participants across nine datasets and found that these factors had as strong an association to cognition as dimensions derived from a supervised approach that expressly maximised the association between each factor and cognitive functioning. Taken together, these findings support the idea that the HiTOP fear and distress subfactors capture coherent, mechanistically valid constructs that capture important differences in underlying cognitive and learning processes. Future work should consider extending these approaches to other HiTOP dimensions to better

understand if, and how, structural distinctions within the HiTOP model are represented in underlying computational and biological processes.

Establishing the Clinical Utility of Dimensions

Data-Driven Approaches to Forecast Clinical Outcomes

A key principle of the HiTOP framework is that a nosology aligned with the structure of psychopathology should enhance clinical practice and outcomes (Ruggero et al., 2019). Most previous studies investigating the clinical utility of HiTOP have focused primarily on *a priori* hypotheses targeting a limited number of dimensions; computational approaches are often unnecessary in these instances, because the ratio of features to outcomes is low and the hypothesis is straightforward. Nonetheless, early evidence for HiTOP's clinical utility is promising: dimensional models outperform traditional diagnoses in predicting clinical improvement, treatment response, and functional outcomes (Cervin et al., 2021; Conway et al., 2021; Forbush et al., 2018, 2024; E. A. Martin et al., 2021).

Moving forward, data-driven methods may play an increasingly important role as researchers begin to (i) examine the many lower-order dimensions within HiTOP (over 90 subscales in the new measure), (ii) explore novel outcomes with uncertain links to individual differences, and (iii) integrate HiTOP dimensions with other data sources. For example, machine learning algorithms can be used to directly compare whether HiTOP dimensions derived solely from self-report data outperform those derived from multimodal inputs in forecasting outcomes (see Fox et al., 2025 for an example). This may help practitioners decide when, and for whom, it is worthwhile to devote additional time and resources to collect biological or computational measures.

Future data-driven validation efforts will require HiTOP to be available in datasets that include multiple data modalities and key clinical outcomes (e.g., treatment trial, longitudinal cohort studies). Large language models (LLMs) may help to address this challenge by deriving HiTOP dimensions from existing datasets, reducing the need to collect additional data. In two studies, an LLM was able to extract dimensional symptom scores (i.e., RDoC domain scores) from clinical notes; these scores predicted clinical outcomes like hospital admission and length of stay (McCoy & Perlis, 2024a, 2024b). If this approach can be extended to the HiTOP framework, it would offer a scalable and efficient approach means of evaluating the clinical utility of HiTOP dimensions in archival data.

Beyond prediction, data-driven methods also offer a way to test whether HiTOP dimensions represent more effective treatment *targets* than traditional diagnoses. Theoretically, better alignment between HiTOP dimensions and underlying cognitive and biological processes should translate into gains in data-driven forecasting of symptom change. Supporting this idea, one recent study found an unsupervised learning model trained on baseline functional connectivity data significantly predicted treatment-related symptom improvement at 12 weeks in a sample of patients with internalizing psychopathology (Zhang et al., 2025). However, most studies in this area remain limited by small samples, a focus on categorical diagnoses, and lack of external validation, so more work is needed.

Intervening on Cognitive Processes to Promote Dimensional Symptom Change

Though data-driven studies are valuable for relating HiTOP dimensions to clinical outcomes, they typically offer little insight into how to intervene effectively, emphasizing

predictive accuracy over mechanism. Conversely, theory-driven approaches are ideally-suited to characterizing the neurocognitive processes that drive individual differences in symptom presentation and change. By modeling these processes as algorithms composed of key parameters, theory-driven approaches allow us to specify the core components of a dysfunctional process and evaluate how targeted modifications – such as adjusting model inputs through clinical intervention – lead to symptom improvement (Brown et al., 2023; Reiter et al., 2021; Yip et al., 2022). Indeed, prior work shows that learning processes are altered in relation to symptom severity, are predictive of outcomes, and are susceptible to intervention (Brown et al., 2024; Huys, Gölzer, et al., 2016; Weigard, Brislin, et al., 2021). Distinct components of interventions also influence specific computational mechanisms (Norbury et al., 2024), and treatment-related changes in learning processes can mediate symptom improvement (Berndt et al., 2025; Fosco et al., 2017; Queirazza et al., 2019; Vrieze et al., 2013).

These early findings suggest that theory-driven computational models may be valuable tools for understanding *how* existing treatments change HiTOP dimensions and for designing interventions that can target individual dimensions more precisely. For example, Brown and colleagues (2021) recently examined reward and loss learning in individuals with and without depression, including a subset that completed 12 weeks of cognitive behavioral therapy (CBT). They found that anhedonia was linked to learning deficits during reward processing, with anhedonic individuals failing to integrate immediate rewards into long-term expectations. During loss learning, negative affect was associated with more negative loss evaluations and weaker prediction errors in the

subgenual anterior cingulate[3]. Neither effect was present when simply considering depression diagnoses (as opposed to dimensional scores). HiTOP's structure provides a useful lens by which to contextualize these findings, since negative affect (dysphoria) typically falls under the internalizing spectrum, whereas anhedonia cross-loads onto detachment, a spectrum that has been linked to variability in reward learning via its close association with extraversion (Blain et al., 2021). Importantly, CBT improved symptoms and normalized disrupted learning parameters, suggesting it functions by enhancing reward learning and reducing negative loss evaluations.

Though not originally designed with HiTOP in mind, this study still provides a useful example of how HiTOP and computational psychiatry can form a happy marriage, with HiTOP dimensions (but not diagnoses) mapping to distinct neurocognitive processes, and modeling providing novel insights into the underlying learning processes that are changing as dimensions improve with treatment. While CBT appears to function as a blunt tool, normalizing both reward and loss learning, this work suggests the possibility of developing interventions that could specifically target one mechanism or the other, potentially yielding more personalized approaches to treatment based on an individual patient's scores across the internalizing and detachment spectra (e.g., Norbury et al., 2024).

Challenges to Integration

Integrating HiTOP and computational psychiatry has immense potential to advance clinical science, but doing so will require overcoming significant practical, methodological, and conceptual challenges. First, there will be steep learning curves for

[3] One caveat to this study is that findings were restricted to the depressed group and did not generalize to healthy controls. This may have been a function of especially strict exclusionary criteria, as controls were required to have no history of any previous psychiatric diagnosis.

researchers in both fields. Many computational researchers will be unfamiliar with HiTOP dimensions (particularly lower-order constructs), the techniques used to derive and measure them, or with the nuances of construct validation. Likewise, many HiTOP researchers have limited exposure to the diverse softwares, coding languages, and modeling frameworks that underpin modern computational methods. Bridging these disciplinary divides will require increased collaboration and cross-training from researchers in both fields.

Additional challenges may be methodological, stemming from the complexities of trying to establish links between constructs at fundamentally different levels of analysis. Indeed, cross-modality effect sizes are often modest in size and require very large sample sizes to detect (Joyner & Perkins, 2023; Marek et al., 2020). One contributing factor to this is the often poor reliability of many behavioral, biological, and computational measures (Elliott et al., 2020; Hedge et al., 2018; Karvelis et al., 2023), which are typically designed to produce robust group-level effects while minimizing between-person individual differences. This design choice inadvertently attenuates the very variability that clinical researchers are most interested in explaining. Moving forward, progress will depend on the field identifying ways to preserve meaningful between-subjects variance in lower-level measures, either by improving task design (Burgoyne et al., 2023; Zorowitz et al., 2023) or adopting better analytic approaches (Bolger et al., 2019; G. Chen et al., 2021; DeYoung, Hilger, et al., 2024; Rouder & Haaf, 2019). Notably, increasing sample size is not a cureall for poor measurement precision (Gell et al., 2024).

Finally, while computational approaches can provide critical external validity evidence for the HiTOP model, less is known about whether this evidence can be used to inform or revise the actual structure of the model. There is a strong desire, both within HiTOP and within computational and biological psychiatry, for our nosology to extend beyond self-report and incorporate distinct data modalities (Huys, Maia, et al., 2016; Joyner & Perkins, 2023; Montague et al., 2012; Ringwald et al., 2025; Wittchen & Beesdo-Baum, 2018). Yet, there is presently no roadmap for how this should unfold. Should the HiTOP model be amended if analyses that integrate biological data with symptom reports reveal dimensions that diverge from the existing HiTOP structure? How should biological or computational indicators be weighed relative to clinical measures? Are there dimensions for which self-report is sufficient to predict important outcomes, and others where including cognitive or biological indicators is distinctly informative? Importantly, the consortium's recently released revisions framework (Forbes, Ringwald, et al., 2024) provides an important foundation for resolving these questions, outlining a transparent process by which the model can evolve. Nonetheless, additional thoughtfulness and care about how neural, behavioral, and computational indices can be integrated will be necessary as multimodal evidence grows and HiTOP begins to wrestle with revisions to the model.

Conclusions

For decades, psychopathology research has been constrained by an overreliance on a flawed descriptive foundation and by tools too imprecise to capture the complexity of mental disorders or their relationship with key outcomes. The emergence of HiTOP and computational psychiatry offers a transformative opportunity

to address these longstanding limitations. Computational psychiatry offers powerful new methods for modeling cognitive and biological processes, but lacks reliable clinical targets and a principled framework for validating its constructs. Conversely, HiTOP provides an empirically grounded system for organizing the core dimensions of psychopathology, but it currently offers limited insight into the mechanisms that produce them. By integrating HiTOP's robust descriptive structure with the mechanistic precision of computational psychiatry, the field can move toward a more unified science of mental illness – one with the potential to enhance diagnosis, guide personalized interventions, and ultimately improve the lives of individuals struggling with psychopathology.

References

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

Allen, T. A., Hall, N. T., Schreiber, A. M., & Hallquist, M. N. (2022). Explanatory personality science in the neuroimaging era: The map is not the territory. *Current Opinion in Behavioral Sciences*, *43*, 236–241.

Allen, T. A., Hallquist, M. N., & Dombrovski, A. Y. (2024). Callousness, exploitativeness, and tracking of cooperation incentives in the human default network. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(29), e2307221121. https://doi.org/10.1073/PNAS.2307221121/SUPPL_FILE/PNAS.2307221121.SAPP.PDF

Allen, T. A., Rueter, A. R., Abram, S. V., Brown, J. S., & DeYoung, C. G. (2017). Personality and neural correlates of mentalizing ability. *European Journal of Personality*, *31*(6). https://doi.org/10.1002/per.2133

Allen, T. A., Schreiber, A. M., Hall, N. T., & Hallquist, M. N. (2020). From Description to Explanation: Integrating Across Multiple Levels of Analysis to Inform Neuroscientific Accounts of Dimensional Personality Pathology. *Journal of Personality Disorders*, *34*(5), 650–676. https://doi.org/10.1521/pedi.2020.34.5.650

Balling, C. E., South, S. C., Lynam, D. R., & Samuel, D. B. (2023). Clinician Perception of the Clinical Utility of the Hierarchical Taxonomy of Psychopathology (HiTOP)

System. *Clinical Psychological Science*, *11*(6), 1108–1121. https://doi.org/10.1177/21677026221138818

Bennett, D., Silverstein, S. M., & Niv, Y. (2019). The two cultures of computational psychiatry. *JAMA Psychiatry*, *76*(6), 563–564. https://doi.org/10.1001/jamapsychiatry.2019.0231

Berndt, L. C., Crawley, D., Tymchyk, L., Hall, A., Onysk, J., Erdmann, T., Wimmer, E., Berwian, I. M., Norbury, A., & Huys, Q. J. (2025). *Early reduction in aversive Pavlovian bias mediates anhedonia improve-ment during Behavioural Activation in realistic treatment settings*.

Blain, S. D., Kent, J. S., Allen, T. A., Lasagna, C. A., Bord, C. A. P. de, Udochi, A. L., Sponheim, S. R., DeYoung, C. G., & Tso, I. F. (2025). Constructs across a hierarchical, dimensional model of psychopathology show differential associations with social and general cognitive ability. *PLOS ONE*, *20*(1), e0317377. https://doi.org/10.1371/journal.pone.0317377

Blain, S. D., Sassenberg, T. A., Xi, M., Zhao, D., & DeYoung, C. G. (2021). Extraversion but not depression predicts reward sensitivity: Revisiting the measurement of anhedonic phenotypes. *Journal of Personality and Social Psychology*, *121*(2), e1–e18. https://doi.org/10.1037/pspp0000371

Blashfield, R. K., & Draguns, J. G. (1976). Evaluative criteria for psychiatric classification. *Journal of Abnormal Psychology*, *85*(2), 140–150. https://doi.org/10.1037/0021-843X.85.2.140

Bolger, N., Zee, K. S., Rossignac-Milon, M., & Hassin, R. R. (2019). Causal processes in psychology are heterogeneous. *Journal of Experimental Psychology: General*, *148*(4), 601–618. https://doi.org/10.1037/xge0000558

Brandl, F., Weise, B., Mulej Bratec, S., Jassim, N., Hoffmann Ayala, D., Bertram, T., Ploner, M., & Sorg, C. (2022). Common and specific large-scale brain changes in major depressive disorder, anxiety disorders, and chronic pain: A transdiagnostic multimodal meta-analysis of structural and functional MRI studies. *Neuropsychopharmacology*, *47*(5), 1071–1080. https://doi.org/10.1038/s41386-022-01271-y

Brown, V. M., Lee, J., Wang, J., Casas, B., & Chiu, P. H. (2024). Reinforcement-Learning-Informed Queries Guide Behavioral Change. *Clinical Psychological Science*, *12*(6), 1146–1161. https://doi.org/10.1177/21677026231213368

Brown, V. M., Price, R., & Dombrovski, A. Y. (2023). Anxiety as a disorder of uncertainty: Implications for understanding maladaptive anxiety, anxious avoidance, and exposure therapy. *Cognitive, Affective, & Behavioral Neuroscience*, *23*(3), 844–868. https://doi.org/10.3758/s13415-023-01080-w

Brown, V. M., Zhu, L., Solway, A., Wang, J. M., McCurry, K. L., King-Casas, B., & Chiu, P. H. (2021). Reinforcement Learning Disruptions in Individuals With Depression and Sensitivity to Symptom Change Following Cognitive Behavioral Therapy. *JAMA Psychiatry*, *78*(10), 1113–1122. https://doi.org/10.1001/jamapsychiatry.2021.1844

Burgoyne, A. P., Tsukahara, J. S., Mashburn, C. A., Pak, R., & Engle, R. W. (2023). Nature and measurement of attention control. *Journal of Experimental Psychology. General*, *152*(8), 2369–2402. https://doi.org/10.1037/xge0001408

Cao, T., Pang, J. C., Gajwani, M., Segal, A., Holmes, A., Wiley, J., Chopra, S., Zhou, J. H., Chen, C. C., Ji, F., Harrison, B. J., Davey, C. G., Constable, T., Tiego, J., Hartshorn, B., Kwee, J., Bellgrove, M. A., & Fornito, A. (2025). *Are neuroanatomical phenotypes for psychiatric disorders robust? An assessment of the reproducibility of grey matter differences in mental illness* (p. 2025.07.09.25331220). medRxiv. https://doi.org/10.1101/2025.07.09.25331220

Carcone, D., Tokarz, V. L., & Ruocco, A. C. (2015). A systematic review on the reliability and validity of semistructured diagnostic interviews for borderline personality disorder. *Canadian Psychology / Psychologie Canadienne*, *56*(2), 208–226. https://doi.org/10.1037/cap0000026

Caspi, A., Houts, R. M., Ambler, A., Danese, A., Elliott, M. L., Hariri, A., Harrington, H., Hogan, S., Poulton, R., Ramrakha, S., Rasmussen, L. J. H., Reuben, A., Richmond-Rakerd, L., Sugden, K., Wertz, J., Williams, B. S., & Moffitt, T. E. (2020). Longitudinal Assessment of Mental Health Disorders and Comorbidities Across 4 Decades Among Participants in the Dunedin Birth Cohort Study. *JAMA Network Open*, *3*(4), e203221. https://doi.org/10.1001/jamanetworkopen.2020.3221

Cervin, M., Norris, L. A., Ginsburg, G., Gosch, E. A., Compton, S. N., Piacentini, J., Albano, A. M., Sakolsky, D., Birmaher, B., Keeton, C., Storch, E. A., & Kendall, P. C. (2021). The *p* Factor Consistently Predicts Long-Term Psychiatric and

Functional Outcomes in Anxiety-Disordered Youth. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(7), 902-912.e5. https://doi.org/10.1016/j.jaac.2020.08.440

Chapman, L. J., & Chapman, J. P. (1978). The measurement of differential deficit. *Journal of Psychiatric Research*, *14*(1–4), 303–311. https://doi.org/10.1016/0022-3956(78)90034-1

Chase, H. W., Frank, M. J., Michael, A., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2010). Approach and avoidance learning in patients with major depression and healthy controls: Relation to anhedonia. *Psychological Medicine*, *40*(3), 433–440. https://doi.org/10.1017/S0033291709990468

Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, *20*(2), 154–170. https://doi.org/10.1002/wps.20882

Chen, G., Pine, D. S., Brotman, M. A., Smith, A. R., Cox, R. W., & Haller, S. P. (2021). Trial and error: A hierarchical modeling approach to test-retest reliability. *NeuroImage*, *245*, 118647. https://doi.org/10.1016/j.neuroimage.2021.118647

Chen, J., Patil, K. R., Yeo, B. T. T., & Eickhoff, S. B. (2023). Leveraging Machine Learning for Gaining Neurobiological and Nosological Insights in Psychiatric Research. *Biological Psychiatry*, *93*(1), 18–28. https://doi.org/10.1016/j.biopsych.2022.07.025

Chen, J., Tam, A., Kebets, V., Orban, C., Ooi, L. Q. R., Asplund, C. L., Marek, S., Dosenbach, N. U. F., Eickhoff, S. B., Bzdok, D., Holmes, A. J., & Yeo, B. T. T.

(2022). Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nature Communications*, *13*(1), 2217. https://doi.org/10.1038/s41467-022-29766-8

Cicchetti, D. (1993). Developmental Psychopathology: Reactions, Reflections, Projections. *Developmental Review*, *13*(4), 471–502. https://doi.org/10.1006/drev.1993.1021

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest*, *18*(2), 72–145. https://doi.org/10.1177/1529100617727266

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*(3), 316–336. https://doi.org/10.1037/0021-843X.100.3.316

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

Conway, C. C., & Brown, T. A. (2018). Evaluating dimensional models of psychopathology in outpatients diagnosed with emotional disorders: A cautionary tale. *Depression and Anxiety*, *35*(9), 898–902. https://doi.org/10.1002/da.22740

Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., Mullins-Sweatt, S. N., Shackman, A. J., Skodol, A. E., South, S. C., Sunderland, M., Waszczuk, M. A., Zald, D. H., Afzali, M. H., Bornovalova, M. A., Carragher,

N., Docherty, A. R., Jonas, K. G., Krueger, R. F., … Eaton, N. R. (2019). A Hierarchical Taxonomy of Psychopathology can transform mental health research. *Perspectives on Psychological Science*, *14*(3), 419–436. https://doi.org/10.1177/1745691618810696

Conway, C. C., Snorrason, I., Beard, C., Forgeard, M., Cuthbert, K., & Björgvinsson, T. (2021). A Higher Order Internalizing Dimension Predicts Response to Partial Hospitalization Treatment. *Clinical Psychological Science*, *9*(3), 373–384. https://doi.org/10.1177/2167702620959291

Costa, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI). *Psychological Assessment Resources*. http://www.getcited.org/pub/103127499

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302.

Cuthbert, B., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, *11*(1), 1–8. https://doi.org/10.1186/1741-7015-11-126/TABLES/2

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *The Behavioral and Brain Sciences*, *22*(3), 491–517.

Depue, R. A., & Lenzenweger, M. F. (2015). Toward a developmental psychopathology of personality disturbances: A neurobehavioral dimensional model. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology (Vol. 2): Developmental neuroscience* (pp. 762–796). Wiley.

Depue, R. A., Luciana, M., Arbisi, P., Collins, P., & Leon, A. (1994). Dopamine and the structure of personality: Relation of agonist-induced dopamine activity to positive emotionality. *Journal of Personality and Social Psychology*, *67*(3), 485–498. https://doi.org/10.1037/0022-3514.67.3.485

DeYoung, C. G., Blain, S. D., Latzman, R. D., Grazioplene, R. G., Haltigan, J. D., Kotov, R., Michelini, G., Venables, N. C., Docherty, A. R., Goghari, V. M., Kallen, A. M., Martin, E. A., Palumbo, I. M., Patrick, C. J., Perkins, E. R., Shackman, A. J., Snyder, M. E., & Tobin, K. E. (2024). The hierarchical taxonomy of psychopathology and the search for neurobiological substrates of mental illness: A systematic review and roadmap for future research. *Journal of Psychopathology and Clinical Science*, *133*(8), 697–715. https://doi.org/10.1037/abn0000903

DeYoung, C. G., Hilger, K., Hanson, J. L., Abend, R., Allen, T., Beaty, R., Blain, S. D., Chavez, R. S., Engel, S. A., Ma, F., Fornito, A., Genç, E., Goghari, V., Grazioplene, R. G., Homan, P., Joyner, K., Kaczkurkin, A. N., Latzman, R. D., Martin, E. A., … Wacker, J. (2024). *Beyond Increasing Sample Sizes: Optimizing Effect Sizes in Neuroimaging Research on Individual Differences*. OSF. https://doi.org/10.31219/osf.io/bjn62

DeYoung, C. G., Kotov, R., Krueger, R. F., Cicero, D. C., Conway, C. C., Eaton, N. R., Forbes, M. K., Hallquist, M. N., Jonas, K. G., Latzman, R. D., Rodriguez-Seijas, C., Ruggero, C. J., Simms, L. J., Waldman, I. D., Waszczuk, M. A., Widiger, T. A., & Wright, A. G. C. (2022). Answering Questions About the Hierarchical Taxonomy of Psychopathology (HiTOP): Analogies to Whales and Sharks Miss

the Boat. *Clinical Psychological Science*, *10*(2), 279–284. https://doi.org/10.1177/21677026211049390

DeYoung, C. G., & Krueger, R. F. (2020). To Wish Impossible Things: On the Ontological Status of Latent Variables and the Prospects for Theory in Psychology. *Psychological Inquiry*, *31*(4), 289–296. https://doi.org/10.1080/1047840X.2020.1853462

Dombrovski, A. Y., Clark, L., Siegle, G. J., Butters, M. A., Ichikawa, N., Sahakian, B. J., & Szanto, K. (2010). Reward/Punishment reversal learning in older suicide attempters. *American Journal of Psychiatry*, *167*(6), 699–707. https://doi.org/10.1176/appi.ajp.2009.09030407

Dück, K., Wüllhorst, R., Overmeyer, R., & Endrass, T. (2024). On the effects of impulsivity and compulsivity on neural correlates of model-based performance. *Scientific Reports*, *14*(1), 21057. https://doi.org/10.1038/s41598-024-71692-w

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis: *Psychological Science*. https://doi.org/10.1177/0956797620916786

Elliott, M. L., Romer, A., Knodt, A. R., & Hariri, A. R. (2018). A Connectome-wide Functional Signature of Transdiagnostic Risk for Mental Illness. *Biological Psychiatry*, *84*(6), 452–459. https://doi.org/10.1016/j.biopsych.2018.03.012

Forbes, M. K., Baillie, A., Batterham, P. J., Calear, A., Kotov, R., Krueger, R. F., Markon, K. E., Mewton, L., Pellicano, E., Roberts, M., Rodriguez-Seijas, C., Sunderland,

M., Watson, D., Watts, A. L., Wright, A. G. C., & Anna Clark, L. (2024). Reconstructing Psychopathology: A Data-Driven Reorganization of the Symptoms in the Diagnostic and Statistical Manual of Mental Disorders. *Clinical Psychological Science*, 21677026241268345. https://doi.org/10.1177/21677026241268345

Forbes, M. K., Neo, B., Nezami, O. M., Fried, E. I., Faure, K., Michelsen, B., Twose, M., & Dras, M. (2024). Elemental psychopathology: Distilling constituent symptoms and patterns of repetition in the diagnostic criteria of the DSM-5. *Psychological Medicine*, *54*(5), 886–894. https://doi.org/10.1017/S0033291723002544

Forbes, M. K., Ringwald, W. R., Allen, T., Cicero, D. C., Clark, L. A., DeYoung, C. G., Eaton, N., Kotov, R., Krueger, R. F., Latzman, R. D., Martin, E. A., Naragon-Gainey, K., Ruggero, C. J., Waldman, I. D., Brandes, C., Fried, E. I., Goghari, V. M., Hankin, B., Sperry, S., … Wright, A. G. C. (2024). Principles and Procedures for Revising the Hierarchical Taxonomy of Psychopathology. *Journal of Psychopathology and Clinical Science*, *133*(1), 4–19. https://doi.org/10.1037/ABN0000886

Forbush, K. T., Chen, P.-Y., Hagan, K. E., Chapa, D. A. N., 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., Chen, Y., Chen, P.-Y., Bohrer, B. K., Hagan, K. E., Iverson-Chapa, D. A. N., Christensen Pacella, K. A., Perko, V., Richson, B. N., Johnson Munguia, S.

N., Thomeczek, M. L., Nelson, S. V., Christian, K., Swanson, T. J., & Wildes, J. E. (2024). Integrating “Lumpers” versus “Splitters” Perspectives: Toward a Hierarchical Dimensional Taxonomy of Eating Disorders from Clinician Ratings. *Clinical Psychological Science*, *12*(4), 625–643. https://doi.org/10.1177/21677026231186803

Fosco, W. D., White, C. N., & Hawk, L. W. (2017). Acute Stimulant Treatment and Reinforcement Increase the Speed of Information Accumulation in Children with ADHD. *Journal of Abnormal Child Psychology*, *45*(5), 911–920. https://doi.org/10.1007/s10802-016-0222-0

Fox, C. A., Teckentrup, V., Donegan, K. R., Seow, T. X., Benwell, C. S., Tervo-Clemmens, B., & Gillan, C. M. (2025). Cognitive arbitration between candidate dimensions of psychopathology. *Molecular Psychiatry*, 1–14. https://doi.org/10.1038/s41380-025-03297-2

Frenay, B., & Verleysen, M. (2014). Classification in the Presence of Label Noise: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(5), 845–869. https://doi.org/10.1109/TNNLS.2013.2292894

Garbusow, M., Ebrahimi, C., Riemerschmid, C., Daldrup, L., Rothkirch, M., Chen, K., Chen, H., Belanger, M. J., Hentschel, A., Smolka, M. N., Heinz, A., Pilhatsch, M., & Rapp, M. A. (2022). Pavlovian-to-Instrumental Transfer across Mental Disorders: A Review. *Neuropsychobiology*, *81*(5), 418–437. https://doi.org/10.1159/000525579

Gell, M., Eickhoff, S. B., Omidvarnia, A., Küppers, V., Patil, K. R., Satterthwaite, T. D., Müller, V. I., & Langner, R. (2024). How measurement noise limits the accuracy of

brain-behaviour predictions. *Nature Communications*, *15*(1), 10678. https://doi.org/10.1038/s41467-024-54022-6

Gillan, C. M., Kalanthroff, E., Evans, M., Weingarden, H. M., Jacoby, R. J., Gershkovich, M., Snorrason, I., Campeas, R., Cervoni, C., Crimarco, N. C., Sokol, Y., Garnaat, S. L., McLaughlin, N. C. R., Phelps, E. A., Pinto, A., Boisseau, C. L., Wilhelm, S., Daw, N. D., & Simpson, H. B. (2020). Comparison of the Association Between Goal-Directed Planning and Self-reported Compulsivity vs Obsessive-Compulsive Disorder Diagnosis. *JAMA Psychiatry*, *77*(1), 77–85. https://doi.org/10.1001/jamapsychiatry.2019.2998

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goaldirected control. *eLife*, *5*(MARCH2016). https://doi.org/10.7554/ELIFE.11305

Gillan, C. M., Vaghi, M. M., Hezemans, F. H., Grothe, S. van G., Dafflon, J., Brühl, A. B., Savulich, G., & Robbins, T. W. (2021). Experimentally induced and real-world anxiety have no demonstrable effect on goal-directed behaviour. *Psychological Medicine*, *51*(9), 1467–1478. https://doi.org/10.1017/S0033291720000203

Goodkind, M., Eickhoff, S. B., Oathes, D. J., Jiang, Y., Chang, A., Jones-Hagata, L. B., Ortega, B. N., Zaiko, Y. V., Roach, E. L., Korgaonkar, M. S., Grieve, S. M., Galatzer-Levy, I., Fox, P. T., & Etkin, A. (2015). Identification of a Common Neurobiological Substrate for Mental Illness. *JAMA Psychiatry*, *72*(4), 305–315. https://doi.org/10.1001/jamapsychiatry.2014.2206

Grahek, I., Schaller, M., & Tackett, J. L. (2021). Anatomy of a Psychological Theory: Integrating Construct-Validation and Computational-Modeling Methods to

Advance Theorizing. *Perspectives on Psychological Science*, *16*(4), 803–815. https://doi.org/10.1177/1745691620966794

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford.

Guest, O., & Martin, A. E. (2021). How Computational Modeling Can Force Theory Building in Psychological Science. *Perspectives on Psychological Science*, *16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-Processing Behavior in Depressed Participants Relative to Healthy Volunteers: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *77*(12), 1286–1295. https://doi.org/10.1001/jamapsychiatry.2020.2139

Hall, N. T., Schreiber, A. M., Allen, T. A., & Hallquist, M. N. (2021). *Disentangling cognitive processes in externalizing psychopathology using drift diffusion modeling: Antagonism, but not disinhibition, is associated with poor cognitive control*. *89*(5), 970–985. https://doi.org/10.1111/jopy.12628

Hallquist, M. N., Allen, T. A., Okan, A., Paolizzi, S., & Dombrovski, A. Y. (2024). Computational psychiatry and dimensional models of psychopathology: Confluences and conflicts. In R. F. Krueger & C. Conway (Eds.), *Handbook of Dimensional Models of Psychopathology*. Oxford University Press.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hopwood, C. J., Bagby, M., Gralnick, T., Ro, E., Ruggero, C., Mullins-Sweatt, S., Kotov, R., Bach, B., Cicero, D., Krueger, R., Patrick, C., Chmielewski, M., DeYoung, C., Docherty, A., Eaton, N., Forbush, K., Ivanova, M., Latzman, R., Pincus, A., … Zimmermann, J. (2018). *Integrating psychotherapy with the Hierarchical Taxonomy of Psychopathology (HiTOP*. https://doi.org/10.31234/osf.io/jb8z4

Huys, Q. J. M., Gölzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., & Dolan, R. J. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological Medicine*, *46*(5), 1027–1035. https://doi.org/10.1017/S0033291715002597

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, *19*(3), 404–413. https://doi.org/10.1038/nn.4238

Huys, Q. J. M., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 12. https://doi.org/10.1186/2045-5380-3-12

Hyman, S. E. (2007). Can neuroscience be integrated into the DSM-V? *Nature Reviews Neuroscience*, *8*(9), 725–732. https://doi.org/10.1038/nrn2218

Joyner, K. J., & Perkins, E. R. (2023). Challenges and ways forward in bridging units of analysis in clinical psychological science. *Journal of Psychopathology and Clinical Science*, *132*(7), 888–896. https://doi.org/10.1037/abn0000879

Karvelis, P., Paulus, M. P., & Diaconescu, A. O. (2023). Individual differences in computational psychiatry: A review of current challenges. *Neuroscience and Biobehavioral Reviews*, *148*. https://doi.org/10.1016/j.neubiorev.2023.105137

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Kebets, V., Holmes, A. J., Orban, C., Tang, S., Li, J., Sun, N., Kong, R., Poldrack, R. A., & Yeo, B. T. T. (2019). Somatosensory-Motor Dysconnectivity Spans Multiple Transdiagnostic Dimensions of Psychopathology. *Biological Psychiatry*, *86*(10), 779–791. https://doi.org/10.1016/j.biopsych.2019.06.013

Kendell, R., & Jablensky, A. (2003). Distinguishing Between the Validity and Utility of Psychiatric Diagnoses. *American Journal of Psychiatry*, *160*(1), 4–12. https://doi.org/10.1176/appi.ajp.160.1.4

Kent, J. S., Markon, K., & MacDonald III, A. W. (2023). Theories of psychopathology: Introduction to a special section. *Journal of Psychopathology and Clinical Science*, *132*(3), 223–227. https://doi.org/10.1037/abn0000824

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

Kind, A., & Dayan, P. (2024). Two sub-cultures of explanatory computational psychiatry. *Molecular Psychiatry*, *29*(12), 3928–3931. https://doi.org/10.1038/s41380-024-02639-w

Kircanski, K., White, L. K., Tseng, W. L., Wiggins, J. L., Frank, H. R., Sequeira, S., Zhang, S., Abend, R., Towbin, K. E., Stringaris, A., Pine, D. S., Leibenluft, E., & Brotman, M. A. (2018). A latent variable approach to differentiating neural

mechanisms of irritability and anxiety in youth. *JAMA Psychiatry*, *75*(6), 631–639. https://doi.org/10.1001/jamapsychiatry.2018.0468

Kist, J. D., Fraza, C., Savage, H. S., Mulders, P. C. R., Vrijsen, J. N., Collard, Rose. M., Tendolkar, I., van Eijndhoven, P., & Marquand, A. F. (2025). Functional network connectivity deviation is associated with transdiagnostic symptomatology. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2025.09.014

Kotov, R., Cicero, D. C., Conway, C. C., DeYoung, C. G., Dombrovski, A., Eaton, N. R., First, M. B., Forbes, M. K., Hyman, S. E., Jonas, K. G., Krueger, R. F., Latzman, R. D., Li, J. J., Nelson, B. D., Regier, D. A., Rodriguez-Seijas, C., Ruggero, C. J., Simms, L. J., Skodol, A. E., … Wright, A. G. C. (2022). The Hierarchical Taxonomy of Psychopathology (HiTOP) in psychiatric practice and research. *Psychological Medicine*, *52*(9), 1666–1678. https://doi.org/10.1017/S0033291722001301

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*. https://doi.org/10.1146/annurev-clinpsy-081219

Kotov, R., Ruggero, C. J., Krueger, R. F., Watson, D., & Zimmerman, M. (2018). The perils of hierarchical exclusion rules: A further word of caution. *Depression and Anxiety*, *35*(9), 903–904. https://doi.org/10.1002/da.22826

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

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., … HiTOP Utility Workgroup. (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. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, *2*, 111–133. https://doi.org/10.1146/annurev.clinpsy.2.022305.095213

Latzman, R. D., Boysen, S. T., & Schapiro, S. J. (2018). Neuroanatomical correlates of hierarchical personality traits in chimpanzees: Associations with limbic structures. *Personality Neuroscience*, *1*(May), 1–11. https://doi.org/10.1017/pen.2018.1

Latzman, R. D., & DeYoung, C. G. (2020). Using empirically-derived dimensional phenotypes to accelerate clinical neuroscience: The Hierarchical Taxonomy of Psychopathology (HiTOP) framework. *Neuropsychopharmacology*, *45*(7), 1083–1085. https://doi.org/10.1038/s41386-020-0639-6

Latzman, R. D., Freeman, H. D., Schapiro, S. J., & Hopkins, W. D. (2015). The contribution of genetics and early rearing experiences to hierarchical personality dimensions in chimpanzees (Pan troglodytes). *Journal of Personality and Social Psychology*, *109*(5), 889–900. https://doi.org/10.1037/pspp0000040

Latzman, R. D., Hecht, L. K., Freeman, H. D., Schapiro, S. J., & Hopkins, W. D. (2015). Neuroanatomical correlates of personality in chimpanzees (Pan troglodytes): Associations between personality and frontal cortex. *NeuroImage*, *123*, 63–71. https://doi.org/10.1016/j.neuroimage.2015.08.041

Lee, P. H., Feng, Y.-C. A., & Smoller, J. W. (2021). Pleiotropy and Cross-Disorder Genetics Among Psychiatric Disorders. *Biological Psychiatry*, *89*(1), 20–31. https://doi.org/10.1016/j.biopsych.2020.09.026

Lloyd, A., Roiser, J. P., Skeen, S., Freeman, Z., Badalova, A., Agunbiade, A., Busakhwe, C., DeFlorio, C., Marcu, A., Pirie, H., Saleh, R., Snyder, T., Fearon, P., & Viding, E. (2024). Reviewing explore/exploit decision-making as a transdiagnostic target for psychosis, depression, and anxiety. *Cognitive, Affective, & Behavioral Neuroscience*, *24*(5), 793–815. https://doi.org/10.3758/s13415-024-01186-9

Love, B. C. (2015). The Algorithmic Level Is the Bridge Between Computation and Brain. *Topics in Cognitive Science*, *7*(2), 230–242. https://doi.org/10.1111/tops.12131

Lutz, M. C., Kok, R., Verveer, I., Malbec, M., Koot, S., van Lier, P. A. C., & Franken, I. H. A. (2021). Diminished error-related negativity and error positivity in children and adults with externalizing problems and disorders: A meta-analysis on error processing. *Journal of Psychiatry and Neuroscience*, *46*(6), E615–E627. https://doi.org/10.1503/jpn.200031

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. https://doi.org/10.1037/1082-989X.7.1.19

Maddux, J. E., Gosselin, J. T., & Winstead, B. A. (2019). Conceptions of Psychopathology: A Social Constructionist Perspective. In *Psychopathology* (5th ed.). Routledge.

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., Moore, L. A., Conan, G., Uriarte, J., … Dosenbach, N. U. F. (2020). *Towards Reproducible Brain-Wide Association Studies* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.08.21.257758

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*, *137*(5), 856–879. https://doi.org/10.1037/a0023678

Marquand, A. F., Kia, S. M., Zabihi, M., Wolfers, T., Buitelaar, J. K., & Beckmann, C. F. (2019). Conceptualizing mental disorders as deviations from normative functioning. *Molecular Psychiatry*, *24*(10), 1415–1424. https://doi.org/10.1038/s41380-019-0441-1

Marr, D. (1982). *Vision* (W. H. Freeman, Ed.).

Martin, E. A., Jonas, K. G., Lian, W., Foti, D., Donaldson, K. R., Bromet, E. J., & Kotov, R. (2021). Predicting Long-Term Outcomes in First-Admission Psychosis: Does the Hierarchical Taxonomy of Psychopathology Aid DSM in Prognostication? *Schizophrenia Bulletin*, *47*(5), 1331–1341. https://doi.org/10.1093/SCHBUL/SBAB043

Martin, J., Taylor, M. J., & Lichtenstein, P. (2018). Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychological Medicine*, *48*(11), 1759–1774. https://doi.org/10.1017/S0033291717003440

Marusak, H. A., Thomason, M. E., Peters, C., Zundel, C., Elrahal, F., & Rabinak, C. A. (2016). You say ‘prefrontal cortex’ and I say ‘anterior cingulate’: Meta-analysis of spatial overlap in amygdala-to-prefrontal connectivity and internalizing symptomology. *Translational Psychiatry*, *6*(11), e944. https://doi.org/10.1038/tp.2016.218

McCoy, T. H., & Perlis, R. H. (2024a). Characterizing research domain criteria symptoms among psychiatric inpatients using large language models. *Journal of Mood & Anxiety Disorders*, *8*, 100079. https://doi.org/10.1016/j.xjmad.2024.100079

McCoy, T. H., & Perlis, R. H. (2024b). Dimensional Measures of Psychopathology in Children and Adolescents Using Large Language Models. *Biological Psychiatry*, *96*(12), 940–947. https://doi.org/10.1016/j.biopsych.2024.05.008

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

McTeague, L. M., Huemer, J., Carreon, D. M., Jiang, Y., Eickhoff, S. B., & Etkin, A. (2017). Identification of Common Neural Circuit Disruptions in Cognitive Control Across Psychiatric Disorders. *American Journal of Psychiatry*, *174*(7), 676–685. https://doi.org/10.1176/appi.ajp.2017.16040400

McTeague, L. M., Rosenberg, B. M., Lopez, J. W., Carreon, D. M., Huemer, J., Jiang, Y., Chick, C. F., Eickhoff, S. B., & Etkin, A. (2020). Identification of Common Neural Circuit Disruptions in Emotional Processing Across Psychiatric Disorders.

American Journal of Psychiatry, *177*(5), 411–421. https://doi.org/10.1176/appi.ajp.2019.18111271

Michelini, G., Carlisi, C. O., Eaton, N. R., Elison, J. T., Haltigan, J. D., Kotov, R., Krueger, R. F., Latzman, R. D., Li, J. J., Levin-Aspenson, H. F., Salum, G. A., South, S. C., Stanton, K., Waldman, I. D., & Wilson, S. (2024). Where do neurodevelopmental conditions fit in transdiagnostic psychiatric frameworks? Incorporating a new neurodevelopmental spectrum. *World Psychiatry*, *23*(3), 333–357. https://doi.org/10.1002/wps.21225

Millner, A. J., Robinaugh, D. J., & Nock, M. K. (2020). Advancing the Understanding of Suicide: The Need for Formal Theory and Rigorous Descriptive Research. *Trends in Cognitive Sciences*, *24*(9), 704–716. https://doi.org/10.1016/j.tics.2020.06.007

Miyake, A., Emerson, M. J., & Friedman, N. P. (2000). Assessment of executive functions in clinical settings: Problems and recommendations. *Seminars in Speech and Language*, *Volume 21*(Number 02), 0169–0183. https://doi.org/10.1055/s-2000-7563

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Mukherjee, D., Filipowicz, A. L. S., Vo, K., Satterthwaite, T. D., & Kable, J. W. (2020). Reward and punishment reversal-learning in major depressive disorder. *Journal of Abnormal Psychology*, *129*(8), 810–823. https://doi.org/10.1037/abn0000641

Nestler, E. J., & Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. *Nature Neuroscience*, *13*(10), 1161–1169. https://doi.org/10.1038/nn.2647

Nikolaidis, A., Chen, A. A., He, X., Shinohara, R., Vogelstein, J., Milham, M., & Shou, H. (2022). *Suboptimal phenotypic reliability impedes reproducible human neuroscience* (p. 2022.07.22.501193). bioRxiv. https://doi.org/10.1101/2022.07.22.501193

Norbury, A., Hauser, T. U., Fleming, S. M., Dolan, R. J., & Huys, Q. J. M. (2024). Different components of cognitive-behavioral therapy affect specific cognitive mechanisms. *Science Advances*, *10*(13), eadk3222. https://doi.org/10.1126/sciadv.adk3222

Olshan, S. M., Richier, C. J., Baacke, K. A., Miller, G. A., & Heller, W. (2024). Making the most of errors: Utilizing erroneous classifications generated by machine-learning models of neuroimaging data to capture disorder heterogeneity. *Journal of Psychopathology and Clinical Science*, *133*(8), 678–689. https://doi.org/10.1037/abn0000943

Panksepp, J. (2004). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford University Press.

Panksepp, J., Fuchs, T., & Iacobucci, P. (2011). The basic neuroscience of emotional experiences in mammals: The case of subcortical FEAR circuitry and implications for clinical anxiety. *Applied Animal Behaviour Science*, *129*(1), 1–17. https://doi.org/10.1016/j.applanim.2010.09.014

Parkes, L., Moore, T. M., Calkins, M. E., Cook, P. A., Cieslak, M., Roalf, D. R., Wolf, D. H., Gur, R. C., Gur, R. E., Satterthwaite, T. D., & Bassett, D. S. (2021).

Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Translational Psychiatry*, *11*(1), 232. https://doi.org/10.1038/s41398-021-01342-6

Pasion, R., & Barbosa, F. (2019). ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. *Neuroscience & Biobehavioral Reviews*, *103*, 133–149. https://doi.org/10.1016/j.neubiorev.2019.06.013

Pasion, R., Ribes-Guardiola, P., Patrick, C., Stewart, R. A., Paiva, T. O., Macedo, I., Barbosa, F., Brislin, S. J., Martin, E. A., Blain, S. D., Cooper, S. E., Ruocco, A. C., Tiego, J., Wilson, S., & Goghari, V. M. (2023). Modeling relations between event-related potential factors and broader versus narrower dimensions of externalizing psychopathology. *Journal of Psychopathology and Clinical Science*, *132*(7), 867–880. https://doi.org/10.1037/abn0000856

Pavlova, B., & Uher, R. (2020). Assessment of Psychopathology: Is Asking Questions Good Enough? *JAMA Psychiatry*, *77*(6), 557–558. https://doi.org/10.1001/jamapsychiatry.2020.0108

Perkins, E. R., Harper, J., Schaefer, J. D., Malone, S. M., Iacono, W. G., Wilson, S., & Patrick, C. J. (2024). Clarifying the place of p300 in the empirical structure of psychopathology over development. *Journal of Psychopathology and Clinical Science*, *133*(8), 733–744. https://doi.org/10.1037/abn0000937

Pike, A. C., & Robinson, O. J. (2022). Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals: A Systematic Review and

Meta-analysis. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2022.0051

Pizzagalli, D. A. (2014). Depression, Stress, and Anhedonia: Toward a Synthesis and Integrated Model. *Annual Review of Clinical Psychology*, *10*(Volume 10, 2014), 393–423. https://doi.org/10.1146/annurev-clinpsy-050212-185606

Preacher, K. J., MacCallum, R. C., Rucker, D. D., & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods*, *10*(2), 178–192. https://doi.org/10.1037/1082-989X.10.2.178

Qu, Y. L., Chen, J., Tam, A., Ooi, L. Q. R., Dhamala, E., Cocuzza, C. V., Zhang, S., Zeng, T., Lawhead, C., Yeo, B. T. T., & Holmes, A. J. (2025). Distinct brain network features predict internalizing and externalizing traits in children, adolescents and adults. *Nature Mental Health*, 1–12. https://doi.org/10.1038/s44220-025-00388-5

Queirazza, F., Fouragnan, E., Steele, J. D., Cavanagh, J., & Philiastides, M. G. (2019). Neural correlates of weighted reward prediction error during reinforcement learning classify response to cognitive behavioral therapy in depression. *Science Advances*, *5*(7), eaav4962. https://doi.org/10.1126/sciadv.aav4962

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

Reininghaus, U., Bohnke, J. R., Chavez-Baldini, U., Gibbons, R., Iveleva, E., Clementz, B. A., Pearlson, Godfrey D., Keshavan, M. S., Sweeney, J. A., & Tamminga, C. A. (2019). *Transdiagnostic dimensions of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP)*. *18*(1), 67–76. https://doi.org/10.1002/wps.20607

Reiter, A. M., Atiya, N. A., Berwian, I. M., & Huys, Q. J. (2021). Neuro-cognitive processes as mediators of psychological treatment effects. *Current Opinion in Behavioral Sciences*, *38*, 103–109. https://doi.org/10.1016/j.cobeha.2021.02.007

Repple, J., Gruber, M., Mauritz, M., de Lange, S. C., Winter, N. R., Opel, N., Goltermann, J., Meinert, S., Grotegerd, D., Leehr, E. J., Enneking, V., Borgers, T., Klug, M., Lemke, H., Waltemate, L., Thiel, K., Winter, A., Breuer, F., Grumbach, P., … Dannlowski, U. (2023). Shared and Specific Patterns of Structural Brain Connectivity Across Affective and Psychotic Disorders. *Biological Psychiatry*, *93*(2), 178–186. https://doi.org/10.1016/j.biopsych.2022.05.031

Ringwald, W. R., Abramovitch, A., Agelink van Rentergem, J. A., & Kotov, R. (2025). Do Cognitive Functions Belong in the Hierarchical Taxonomy of Psychopathology Model? A Meta-Analysis. *Perspectives on Psychological Science*, 17456916251347926. https://doi.org/10.1177/17456916251347926

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*, 533–546. https://doi.org/10.1017/S0033291721001902

Romer, A. L., Knodt, A. R., Sison, M. L., Ireland, D., Houts, R., Ramrakha, S., Poulton, R., Keenan, R., Melzer, T. R., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2021). Replicability of structural brain alterations associated with general psychopathology: Evidence from a population-representative birth cohort. *Molecular Psychiatry*, *26*(8), 3839–3846. https://doi.org/10.1038/s41380-019-0621-z

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Royer, J., Kebets, V., Piguet, C., Chen, J., Ooi, L. Q. R., Kirschner, M., Siffredi, V., Misic, B., Yeo, B. T., & Bernhardt, B. C. (2024). Multimodal neural correlates of childhood psychopathology. *eLife*, *13*, e87992. https://doi.org/10.7554/eLife.87992

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E., Mullins-Sweatt, S. N., Patrick, C. J., Bach, B., Cicero, D. C., Docherty, A., Simms, L. J., Michael Bagby, R., Krueger, R. F., Callahan, J. L., Chmielewski, M., Conway, C. C., Clercq, B. D., Dornbach-Bender, A., … Zimmermann, J. (2019). Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into Clinical Practice. *Journal of Consulting and Clinical Psychology*, *87*(12), 1069. https://doi.org/10.1037/CCP0000452

Ruscio, A. M. (2019). Normal Versus Pathological Mood: Implications for Diagnosis. *Annual Review of Clinical Psychology*, *15*, 179–205. https://doi.org/10.1146/annurev-clinpsy-050718-095644

Scott, I., & Nelson, B. (2024). Are Psychiatric Nosologies Limiting the Success of Clinical Prediction Models? *JAMA Psychiatry*, *81*(11), 1055–1056. https://doi.org/10.1001/jamapsychiatry.2024.2662

Segal, A., Parkes, L., Aquino, K., Kia, S. M., Wolfers, T., Franke, B., Hoogman, M., Beckmann, C. F., Westlye, L. T., Andreassen, O. A., Zalesky, A., Harrison, B. J., Davey, C. G., Soriano-Mas, C., Cardoner, N., Tiego, J., Yücel, M., Braganza, L., Suo, C., … Fornito, A. (2023). Regional, circuit and network heterogeneity of brain abnormalities in psychiatric disorders. *Nature Neuroscience*, *26*(9), 1613–1629. https://doi.org/10.1038/s41593-023-01404-6

Seow, T. X. F., Benoit, E., Dempsey, C., Jennings, M., Maxwell, A., O'Connell, R., & Gillan, C. M. (2021). Model-Based Planning Deficits in Compulsivity Are Linked to Faulty Neural Representations of Task Structure. *Journal of Neuroscience*, *41*(30), 6539–6550. https://doi.org/10.1523/JNEUROSCI.0031-21.2021

Shackman, A. J., Stockbridge, M. D., Tillman, R. M., Kaplan, C. M., Tromp, D. P. M., Fox, A. S., & Gamer, M. (2016). The neurobiology of dispositional negativity and attentional biases to threat: Implications for understanding anxiety disorders in adults and youth. *Journal of Experimental Psychopathology*, *7*(3), 311–342. https://doi.org/10.5127/jep.054015

Sharma, A., Wolf, D. H., Ciric, R., Kable, J. W., Moore, T. M., Vandekar, S. N., Katchmar, N., Daldal, A., Ruparel, K., Davatzikos, C., Elliott, M. A., Calkins, M. E., Shinohara, R. T., Bassett, D. S., & Satterthwaite, T. D. (2017). Common Dimensional Reward Deficits Across Mood and Psychotic Disorders: A

Connectome-Wide Association Study. *American Journal of Psychiatry*, *174*(7), 657–666. https://doi.org/10.1176/appi.ajp.2016.16070774

Sheffield, J. M., Kandala, S., Tamminga, C. A., Pearlson, G. D., Keshavan, M. S., Sweeney, J. A., Clementz, B. A., Lerman-Sinkoff, D. B., Hill, S. K., & Barch, D. M. (2017). Transdiagnostic Associations Between Functional Brain Network Integrity and Cognition. *JAMA Psychiatry*, *74*(6), 605–613. https://doi.org/10.1001/jamapsychiatry.2017.0669

Simms, L. J. (2023, April 1). *Introduction of Self-Report Measures of the HiTOP Model: Reliability and Structural Validity of the HiTOP-SR*. 2023 SPA Convention and Workshops. https://personality.confex.com/personality/2023/meetingapp.cgi/Paper/1658

Smaldino, P. E. (2017). Models Are Stupid, and We Need More of Them. In *Computational Social Psychology*. Routledge.

Smaldino, P. E. (2020). How to Translate a Verbal Theory Into a Formal Model. *Https://Doi.Org/10.1027/1864-9335/A000425*, *51*(4), 207–218. https://doi.org/10.1027/1864-9335/A000425

Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., Khalsa, S. S., Feinstein, J., Paulus, M. P., & Aupperle, R. L. (2021). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: A computational modelling approach. *Journal of Psychiatry and Neuroscience*, *46*(1), E74–E87. https://doi.org/10.1503/jpn.200032

Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2019). Psychiatric genetics and the structure of psychopathology.

Molecular Psychiatry, *24*(3), 409–420. https://doi.org/10.1038/s41380-017-0010-4

Sripada, C., & Weigard, A. (2021). Impaired Evidence Accumulation as a Transdiagnostic Vulnerability Factor in Psychopathology. *Frontiers in Psychiatry*, *12*, 627179. https://doi.org/10.3389/FPSYT.2021.627179

Sroufe, L. A. (1990). Considering normal and abnormal together: The essence of developmental psychopathology. *Development and Psychopathology*, *2*(4), 335–347. https://doi.org/10.1017/s0954579400005769

Stephan, K. E., Iglesias, S., Heinzle, J., & Diaconescu, A. O. (2015). Translational Perspectives for Computational Neuroimaging. *Neuron*, *87*(4), 716–732. https://doi.org/10.1016/j.neuron.2015.07.008

Stephan, K. E., & Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology*, *25*, 85–92. https://doi.org/10.1016/J.CONB.2013.12.007

Su, C. C., Clark, L. A., Hopwood, C. J., Jonas, K. G., Kotov, R., Nelson, B. D., Rodriguez-Seijas, C., Ruggero, C. J., Simms, L. J., Mullins-Sweatt, S. N., & Cicero, D. C. (2025). Facilitating the implementation of the Hierarchical Taxonomy of Psychopathology (HiTOP) in clinical practice: A feasibility study. *Professional Psychology: Research and Practice*. https://doi.org/10.1037/pro0000644

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press.

Taylor, J. J., Lin, C., Talmasov, D., Ferguson, M. A., Schaper, F. L. W. V. J., Jiang, J., Goodkind, M., Grafman, J., Etkin, A., Siddiqi, S. H., & Fox, M. D. (2023). A

Transdiagnostic Network for Psychiatric Illness Derived from Atrophy and Lesions. *Nature Human Behaviour*, *7*(3), 420–429. https://doi.org/10.1038/s41562-022-01501-9

Tomlinson, R. C., Weigard, A. S., Sripada, C., Jonides, J., Klump, K. L., Burt, S. A., & Hyde, L. W. (2025). Efficiency of evidence accumulation as a formal model-based measure of task-general executive functioning in adolescents. *Journal of Psychopathology and Clinical Science*, *134*(7), 761–774. https://doi.org/10.1037/abn0001043

Trayvick, J., Adams, E. M., & Nelson, B. D. (2025). The P300 and hierarchical dimensions of psychopathology. *Journal of Psychopathology and Clinical Science*, *134*(2), 132–142. https://doi.org/10.1037/abn0000955

Treadway, M. T., & Zald, D. H. (2013). Parsing Anhedonia: Translational Models of Reward-Processing Deficits in Psychopathology. *Current Directions in Psychological Science*, *22*(3), 244–249. https://doi.org/10.1177/0963721412474460

Vanyukov, P. M., Hallquist, M. N., Delgado, M., Szanto, K., & Dombrovski, A. Y. (2019). Neurocomputational mechanisms of adaptive learning in social exchanges. *Cognitive, Affective & Behavioral Neuroscience*, *19*(4), 985–997. https://doi.org/10.3758/s13415-019-00697-0

Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., Schmidt, M., & Claes, S. (2013). Reduced Reward Learning Predicts Outcome in Major Depressive Disorder. *Biological Psychiatry*, *73*(7), 639–645. https://doi.org/10.1016/j.biopsych.2012.10.014

Waldman, I. D., Poore, H. E., Luningham, J. M., & Yang, J. (2020). Testing structural models of psychopathology at the genomic level. *World Psychiatry*, *19*(3), 350–359. https://doi.org/10.1002/wps.20772

Waszczuk, M. A., Eaton, N. R., Krueger, R. F., Shackman, A. J., Waldman, I. D., Zald, D. H., Lahey, B. B., Patrick, C. J., Conway, C. C., Ormel, J., Hyman, S. E., Fried, E. I., Forbes, M. K., Docherty, A. R., Althoff, R. R., Bach, B., Chmielewski, M., DeYoung, C. G., Forbush, K. T., … Kotov, R. (2020). Redefining phenotypes to advance psychiatric genetics: Implications from hierarchical taxonomy of psychopathology. *Journal of Abnormal Psychology*, *129*(2), 143–161. https://doi.org/10.1037/abn0000486

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

Weigard, A., Angstadt, M., Taxali, A., Heathcote, A., Heitzeg, M. M., & Sripada, C. (2024). Flexible adaptation of task-positive brain networks predicts efficiency of evidence accumulation. *Communications Biology*, *7*(1), 801. https://doi.org/10.1038/s42003-024-06506-w

Weigard, A., Brislin, S. J., Cope, L. M., Hardee, J. E., Martz, M. E., Ly, A., Zucker, R. A., Sripada, C., & Heitzeg, M. M. (2021). Evidence accumulation and associated

error-related brain activity as computationally-informed prospective predictors of substance use in emerging adulthood. *Psychopharmacology*, *238*(9), 2629–2644. https://doi.org/10.1007/s00213-021-05885-w

Weigard, A., Clark, D. A., & Sripada, C. (2021). Cognitive efficiency beats top-down control as a reliable individual difference dimension relevant to self-control. *Cognition*, *215*, 104818. https://doi.org/10.1016/J.COGNITION.2021.104818

Weigard, A., & Sripada, C. (2021). Task-General Efficiency of Evidence Accumulation as a Computationally Defined Neurocognitive Trait: Implications for Clinical Neuroscience. *Biological Psychiatry Global Open Science*, *1*(1), 5–15. https://doi.org/10.1016/J.BPSGOS.2021.02.001

Williams, C. M., Peyre, H., Wolfram, T., Lee, Y. H., Seidlitz, J., Ge, T., Smoller, J. W., Mallard, T. T., & Ramus, F. (2024). Characterizing the phenotypic and genetic structure of psychopathology in UK Biobank. *Nature Mental Health*, *2*(8), 960–974. https://doi.org/10.1038/s44220-024-00272-8

Wise, T., & Dolan, R. J. (2020). Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. *Nature Communications*, *11*(1), 1–13. https://doi.org/10.1038/s41467-020-17977-w

Wise, T., Robinson, O. J., & Gillan, C. M. (2023). Identifying Transdiagnostic Mechanisms in Mental Health Using Computational Factor Modeling. *Biological Psychiatry*, *93*(8), 690–703. https://doi.org/10.1016/j.biopsych.2022.09.034

Wittchen, H., & Beesdo‐Baum, K. (2018). “Throwing out the baby with the bathwater”? Conceptual and methodological limitations of the HiTOP approach. *World Psychiatry*, *17*(3), 298–299. https://doi.org/10.1002/wps.20561

Xia, C. H., Ma, Z., Ciric, R., Gu, S., Betzel, R. F., Kaczkurkin, A. N., Calkins, M. E., Cook, P. A., Garza, A. G. de la, Vandekar, S. N., Cui, Z., Moore, T. M., Roalf, D. R., Ruparel, K., Wolf, D. H., Davatzikos, C., Gur, R. C., Gur, R. E., Shinohara, R. T., … Satterthwaite, T. D. (2018). Linked dimensions of psychopathology and connectivity in functional brain networks. *Nature Communications*, *9*(1), 1–14. https://doi.org/10.1038/s41467-018-05317-y

Yarkoni, T. (2015). Neurobiological substrates of personality: A critical overview. In M. Mikulincer & P. R. Shaver (Eds.), *APA handbook of personality and social psychology, Volume 4: Personality processes and individual differences.* (pp. 61–84). American Psychological Association. https://doi.org/10.1037/14343-003

Yip, S. W., Barch, D. M., Chase, H. W., Flagel, S., Huys, Q. J. M., Konova, A. B., Montague, R., & Paulus, M. (2022). From computation to clinic. *Biological Psychiatry Global Open Science*, S2667174322000507. https://doi.org/10.1016/j.bpsgos.2022.03.011

Zhang, K., Klumpp, H., Jimmy, J., Phan, K. L., Milad, M. R., & Wen, Z. (2025). Functional Connectivity Predicting Transdiagnostic Treatment Outcomes in Internalizing Psychopathologies. *JAMA Network Open*, *8*(9), e2530008. https://doi.org/10.1001/jamanetworkopen.2025.30008

Zorowitz, S., Karni, G., Paredes, N., Daw, N., & Niv, Y. (2023). *Improving the reliability of the Pavlovian go/no-go task for computational psychiatry research*. PsyArXiv. https://doi.org/10.31234/osf.io/eb697

Figures

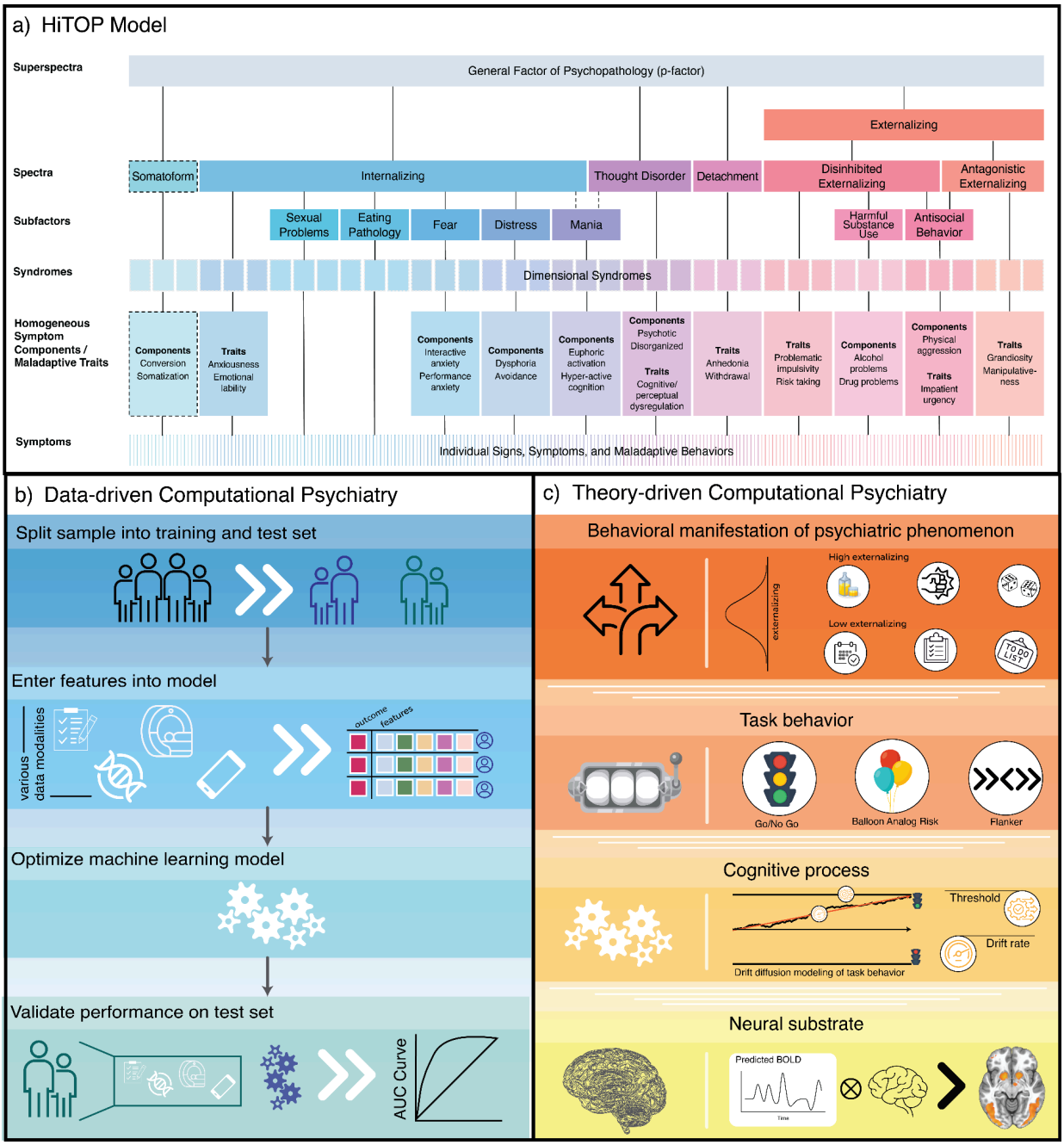

Figure 1. a) Abridged version of current HiTOP model. b) Overview of a data-driven pipeline that combines multimodal data in service of predicting scores on a HiTOP dimension. c) Overview of a theory-driven computational psychiatry approach to studying latent cognitive processes underlying structural dimensions of psychopathology.

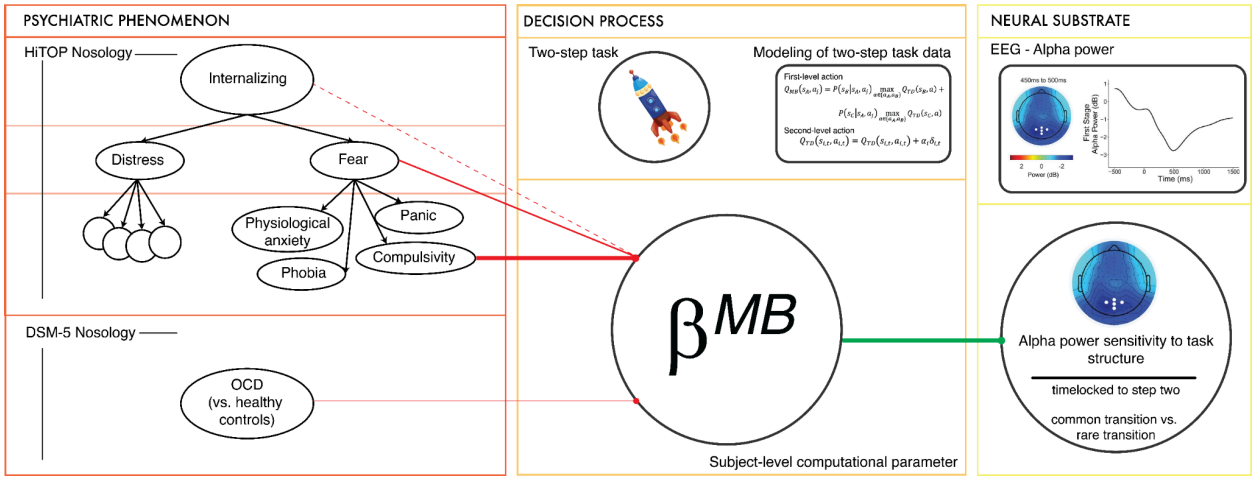

Figure 2. Association between psychiatric phenomenon and computational parameter (β^{MB}) is stronger when examining relevant HiTOP dimension (compulsivity) than when examining group differences between cases (OCD) and controls. Computational parameter (β^{MB}) relates to neural processing (alpha power in EEG). N.B. Green is used to denote a positive association, red to denote a negative association, and line weight to denote strength of association.

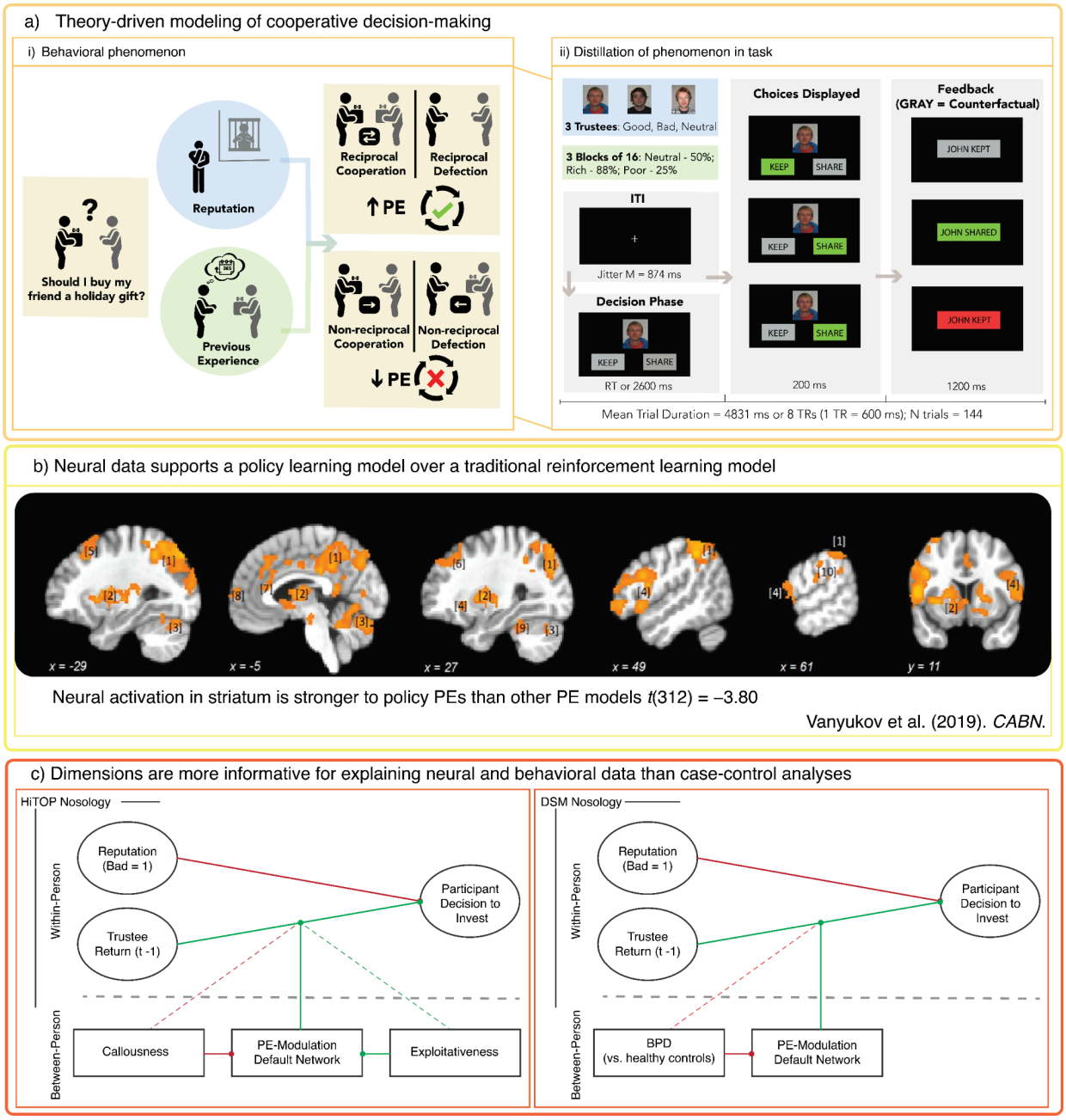

Figure 3. a) In daily life, individuals learn how to interact with others. Rather than solely maximizing rewards, people track the extent to which their policy toward others is correct – (i) such as, whether you correctly predict whether a neighbor will buy you a holiday gift. (ii) Similar learning processes can be captured using a modified trust game. b) Policy vs. reward learning rules make dissociable predictions about behavior and about neural processing. During the trust game, neural dynamics support a policy learning model (compared to classic reward prediction errors; Vanyukov et al., 2019). c) Facets of antagonism – callousness vs. exploitativeness – differentially predict neural processing and decision-making on the trust game. Corresponding group analyses (BPD vs. HC) recapitulate only one of the findings from the primary analyses (Allen et

al., 2024, *PNAS*). Green denotes a positive association. Red denotes a negative association. Dashed line denotes a significant indirect effect.