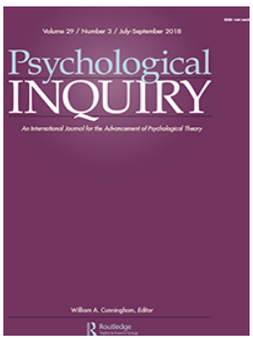

Routledge
Taylor & Francis Group

Psychological Inquiry

An International Journal for the Advancement of Psychological Theory

ISSN: 1047-840X (Print) 1532-7965 (Online) Journal homepage: http://www.tandfonline.com/loi/hpli20

A Cybernetic Theory of Psychopathology

Colin G. DeYoung & Robert F. Krueger

To cite this article: Colin G. DeYoung & Robert F. Krueger (2018) A Cybernetic Theory of Psychopathology, Psychological Inquiry, 29:3, 117-138, DOI: 10.1080/1047840X.2018.1513680

To link to this article: https://doi.org/10.1080/1047840X.2018.1513680

Published online: 24 Dec 2018.

PSYCHOLOGICAL INQUIRY
2018, VOL. 29, NO. 3, 117–138
https://doi.org/10.1080/1047840X.2018.1513680

Routledge
Taylor & Francis Group

TARGET ARTICLE

A Cybernetic Theory of Psychopathology

Colin G. DeYoung and Robert F. Krueger

Department of Psychology, University of Minnesota, Minneapolis, Minnesota

ABSTRACT

Cybernetics, the study of principles governing goal-directed, self-regulating systems, offers a useful approach to understanding psychopathology or psychological dysfunction, overcoming limitations of other naturalistic approaches. Whereas influential theories of psychopathology have relied on definitions of dysfunction rooted in evolution and fitness, we define psychopathology in terms of cybernetic dysfunction, failure to make progress toward important goals. Cybernetic function in organisms is not identical to evolutionary function, despite their causal phylogenetic relationship. We define psychopathology as persistent failure to move toward one's goals, due to failure to generate effective new goals, interpretations, or strategies when existing ones prove unsuccessful. This definition allows a thorough integration of dimensional approaches to psychopathology and personality and provides a new perspective on the nosology of mental disorder. We review evidence that the major dimensions of psychopathology correspond to major trait dimensions of personality, but we assert that extremity on these dimensions is neither necessary nor sufficient for psychopathology, which requires cybernetic dysfunction. Drawing from psychological and neurobiological research on personality and psychopathology, we present a theory of the mechanisms underlying the five major dimensions of psychopathology, some of their subdimensions, and the general risk factor for psychopathology. We conclude by discussing implications of our theory for research, diagnosis, and mental health interventions.

KEYWORDS

Characteristic adaptations; cybernetics; goals; mental disorder; psychopathology; personality; traits

Surprisingly little progress has been made in uncovering the sources of mental disorder. This failure can be attributed in large part to two errors: (a) the characterization of mental disorders as distinct, categorical entities and (b) the assumption that their sources in the brain are best described in terms of one or a few biological abnormalities that are specific to each of the putatively distinct categories (e.g., the dopamine theory of schizophrenia; Kendler & Schaffner, 2011). These errors are encouraged and exacerbated by lack of an adequate understanding of the nature of mental disorder as such. We introduce a theory of psychopathology (psychological dysfunction) that provides a new understanding of mental disorders (diagnostic entities) and corrects the two errors.

In relation to the first error, we rely on a body of evidence showing that most symptoms of mental disorders are on a continuum with normal personality traits (rather than existing as their own distinct categories) and that the symptoms that tend to appear together have a very similar covariance structure to normal personality traits. In relation to the second, we draw on a cybernetic theory of personality that defines dysfunction in terms of inability to move toward important goals: cybernetic Big Five theory (CB5T; DeYoung, 2015). Cybernetics (also known as "control theory") is the study of principles governing goal-directed systems that self-regulate via feedback,[1] and it provides a powerful framework for understanding psychological and brain function (Austin & Vancouver, 1996; Bechhoefer, 2005; Carver & Scheier, 1998; Powers, 1973; Wiener, 1961). Drawing on CB5T allows us to begin identifying the underlying causes of dysfunction in each of the major dimensions of psychopathology, in terms of psychological processes that can be functionally unified but are instantiated by complex networks of brain systems. CB5T also allows us to provide new definitions of *mental disorder* and *psychopathology* that are crucial to making progress in understanding etiology and treatment.

We begin our endeavor with the question of definition. Having made an argument for a cybernetic conception of psychopathology and its relation to personality, we then describe the hierarchical, multidimensional structure of psychopathology and attempt to identify the likely mechanistic sources of that structure. We draw upon evidence indicating that dimensions of variation in psychopathology map onto the Big Five personality traits that constitute the most widely used model of normal personality variation: Neuroticism,

CONTACT Colin G. DeYoung cdeyoung@umn.edu Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455.

[1]Feedforward regulation, in which the system uses information regarding the current state of the world to predict a likely future state and guides action according to that prediction, is also common in complex cybernetic systems, including many organisms (Bechhoefer, 2005). It is certainly involved in most human behavior, in conjunction with feedback processes (Del Giudice, 2015). However, feedforward is not part of the minimal necessary definition of a cybernetic system, whereas feedback is.

Agreeableness, Conscientiousness, Extraversion, and Openness/Intellect (John, Naumann, & Soto, 2008). Psychological and neurobiological research on these trait dimensions complements existing research on mental disorders to help delineate mechanisms of psychopathology.

As a whole, our theory is intended to provide a mechanistic account of psychopathology that can inform both scientific and clinical thinking. It delineates functionally coherent individual differences that help make sense of the various manifestations of psychopathology—that is, of the variety and comorbidity of mental disorders. Further, it offers a scaffold for neurobiological inquiry into the etiology of mental disorder that should be much more effective than current diagnostic categories.

Defining Psychopathology

Mental Disorder versus Psychopathology

Approaches to defining "mental disorder" range from purely value-based accounts, which assert that mental disorders are merely cultural constructions deriving from various sociopolitical concerns, to purely naturalistic accounts, which assert that mental disorders are exclusively a matter of objective fact. In between these extremes are hybrid accounts, involving a combination of value and fact in the determination of disorder. The most influential of the hybrid accounts (and indeed the most influential account of any type) is Wakefield's (1992a, 1992b) *harmful dysfunction* theory, which asserts that "a condition is a disorder if it is negatively valued ('harmful') and it is in fact due to a failure of some internal mechanism to perform a function for which it was biologically designed (i.e., naturally selected)" (Wakefield, 2007, p. 149). "Harmful" is further glossed as "judged negative by sociocultural standards" (Wakefield, 2007, p. 149).

Like Wakefield, we take a hybrid approach, inasmuch as we assert that, although dysfunction is a matter of fact, the diagnosis and treatment of mental disorders in a clinical context requires culturally constructed judgments about the criteria for disorders. Unlike Wakefield, however, we do not assert that harm is entirely culturally constructed. Rather, harm can be defined in relation to the individual rather than socioculturally and is largely overlapping with dysfunction (once dysfunction is properly understood cybernetically). Further, we believe that only the *degree* of dysfunction qualifying as a disorder is inescapably value-laden and culturally constructed. We will focus primarily on *psychopathology*, which we consider synonymous with "psychological dysfunction," rather than on *mental disorder* (officially sanctioned diagnoses), precisely because we do not intend to address the degree of dysfunction necessary for specific diagnoses and interventions. Establishing thresholds for mental disorders is a pragmatic project that must be hashed out in a broader cultural context. Psychopathology becomes mental disorder based on a culturally negotiated judgment that it has passed a level of severity deemed worthy of treatment (and this level will differ for different kinds of psychopathology and in different contexts). We do not believe there is any purely objective standard that can be applied to make such judgments. We do believe, however, that a valid objective standard exists for identifying psychopathology (although, for reasons explained later, its application to individual cases can involve some uncertainty).

Evolutionary versus Cybernetic Dysfunction

Where our theory diverges most importantly from Wakefield's is in defining dysfunction. Many naturalistic approaches to defining mental disorder or psychopathology, including the naturalistic component of hybrid approaches like Wakefield's, rely on an evolutionary definition of function and dysfunction. In such approaches, a mental mechanism's function is that for which it was selected by evolution—that is, the activity that allowed it to increase fitness, successful reproduction over generations (Wakefield, 1992a, 1992b, 2007). Adopting this evolutionary meaning of function introduces a fatal flaw into attempts to specify naturalistic criteria, and indeed, it is this aspect of Wakefield's argument that has been most extensively critiqued. Most critiques of evolutionarily based naturalistic accounts have focused on the inadequacies of the assertions that psychological processes are necessarily dysfunctional if they are not in accord with their cause of selection and, conversely, that processes are not dysfunctional if they are in accord with their cause of selection (e.g., Lilienfeld & Marino, 1995, 1999; Richters & Hinshaw, 1999).[2] We do not reiterate all of these critiques here (we find their general form compelling, even when some of their details can be rebutted; e.g., Wakefield, 1999), but we do address what we believe is the root cause of the problem: Two distinct kinds of function and dysfunction are often not separated conceptually. These are evolutionary and cybernetic.

A mechanism's evolutionary function is its manner of increasing reproductive fitness that caused it to become species-typical. It is not always easy to determine whether and why a particular feature of an organism was evolutionarily selected, but the basic idea is straightforward. This is probably the most common usage for "function" in biology, but there are several others (Wouters, 2003). A mechanism's cybernetic function is its manner of facilitating the goal pursuit of the system of which it is a part. A thorough understanding of this definition requires a description of cybernetic systems and their necessary components.

Cybernetic systems must have three basic components: (a) one or more goals, (b) a representation of the current state of the world in relation to the goal(s), and (c) a set of operators that allow movement toward the goal(s). In cybernetic terms, a goal is a value or range of values of a variable for which the system acts to bring that variable toward or

[2]Another relevant critique is that reliance on evolutionary function to define psychopathology has the consequence that the validity of any mental disorder will depend on a theoretical claim that is extremely difficult to prove, namely, the reasons for the evolution of whatever psychological features of the organism are relevant to the disorder (Bolton, 2008). Our theory of psychopathology as cybernetic dysfunction circumvents this problem because identifying disruption of an individual's goal pursuit is considerably more tractable than determining evolutionary forces.

away from that value or range.[3] Thus, goals involve a controlled variable represented physically in some manner within the system. To take one of the simplest cybernetic mechanisms as an example, the goal of a thermostat is the desired temperature that is set by the user. Pursuit of the goal is possible due to feedback, which requires some way for the system to measure the current state of the controlled variable—a thermometer in the case of the thermostat. When feedback indicates that the current state does not match the goal state, operators are engaged to change the state. In the thermostat, these would be signals that turn on or off heating and cooling systems.

Organisms contain many variables that are controlled cybernetically through feedback, and the very existence of organisms as cybernetic systems is caused by evolution because reproductive fitness is facilitated by the organism's pursuit of various goals. However, as noted by Gray (2004), reproductive fitness is not a variable that is itself physically represented within the organism, and thus fitness is not a goal of the living cybernetic system. (Nor is there any sense in which evolution itself has goals, being merely a probabilistic process, wherein certain features are selected by their reproductive consequences for each generation and hence increase in frequency.) Of course, most goals of living systems have been directly selected by evolution so that their cybernetic and evolutionary functions overlap, but cybernetic and evolutionary functions need not be identical. Particularly for human beings, they are often not, because we have evolved an apparently unprecedented degree of flexibility in the goals we can adopt.

This distinction between cybernetic and evolutionary function has occasionally been confused in the literature relevant to defining psychopathology. For example, one of the most influential naturalistic accounts is by Boorse (1977, 2011), who explicitly describes his account as based in cybernetics but mischaracterizes the cybernetic definition of "goal" by failing to specify that goals must be represented within the cybernetic system. This allows him to argue that "for physiology, the highest level goals, of the organism as a whole, are individual survival and reproduction" (Boorse, 2011, p. 27). Although these two outcomes are crucial for evolutionary function, they are not necessarily the organism's highest level goals, or even (especially for reproduction) goals it possesses at all. It is easier for evolution to select for a goal of engaging in sex than for it to select for a goal of reproduction, for example, and, barring human ingenuity in birth control, only the former is necessary for fitness. There are many examples of human behaviors that are contrary to reproduction (e.g., celibacy) and even survival (e.g., a principled hunger strike) but are not necessarily pathological. Such behaviors are very likely to be evolutionarily dysfunctional, but they are not necessarily cybernetically dysfunctional. Boorse's account, therefore, is rooted in an evolutionary account of function and is not properly considered cybernetic.[4] Note, however, that Boorse explicitly does not define "function" in terms of evolutionary fitness and selection but rather simply as whatever species-typical action supports current survival or reproduction, regardless of why it originally became species-typical. Perhaps this position should be called "quasi-evolutionary" as the specification of survival and reproduction as the most important outcomes for an organism is clearly rooted in evolutionary theory.

We assert that cybernetic dysfunction, not evolutionary dysfunction, is what is relevant to defining psychopathology, and when we refer to "dysfunction" we mean "cybernetic dysfunction" unless otherwise indicated. In addition, we believe that the kind of dysfunction that is typically intended in clinical and lay conceptions of mental disorder resembles cybernetic dysfunction, not evolutionary dysfunction. Although we did not consider clinical or lay usage as a criterion in developing our account, we nonetheless believe that its correspondence with those usages is a benefit. Our position is that whether an evolutionary function (i.e., one that increased fitness in the environment of evolutionary adaptedness) is being fulfilled is irrelevant for psychopathology, except inasmuch as that function is also a cybernetic function governed by a goal of the individual in question. (Note that our avoidance of evolutionary function as a criterion for psychopathology nonetheless leaves open the possibility that analysis of evolutionary function may be useful for understanding the sources of risk for different forms of psychopathology; Del Giudice, 2016).

The Cybernetics of Psychopathology

We believe that a hybrid definition of disorder as *cybernetic dysfunction deemed severe enough for diagnosis and treatment* could probably be extended to all medical pathology, not just psychopathology, but we do not attempt this extension here because most of the details of our theory are specific to psychopathology. This means that the cybernetic goals relevant to our theory are psychological goals. What are psychological goals? Organisms contain cybernetic subsystems regulating many things—temperature, metabolism,

[3]Goals for which the system acts to move away from the value or out of the range of values are avoidance goals. An important asymmetry exists between approach and avoidance goals because, except in moments of sheer panic, avoidance goals are unlikely to govern behavior without a simultaneously operative approach goal. This is because wanting not to do something or to avoid something does not specify what to do or where to go instead (Carver & Scheier, 1998; Mansell, 2005). Formally, activation of approach goals tends to reduce the entropy of the cybernetic system, whereas activation of avoidance goals tends to increase it.

[4]Boorse's account suffers not only from the confusion of evolutionary and cybernetic function but also from an unacknowledged appeal to value. Boorse described his account as value-free, yet his primary criterion for disorder is deviation from statistical normality. Even ignoring the difficulty of establishing a population norm (Boorse acknowledged that different norms are necessary at different ages and for different sexes), this account cannot be value free because of the fact that individual variation is the norm. Indeed, evolution cannot occur without variation in characteristics that have consequences for fitness. This means that one cannot identify disorder with any deviation from the norm (even if, like Boorse, one limits the criterion to deviation in the direction of reduced function) because every individual will vary from the norm in vastly many ways. Thus, Boorse's account requires identifying how much distance from the norm in any given variable is required for identifying disorder, and he himself acknowledged that this distance is arbitrary. Any such arbitrary decision entails a socioculturally negotiated value judgment. Hence, Boorse's account is value-laden and, thus, hybrid (cf. Bolton, 2008).

defense against pathogens, and so on—and each has its own goal(s) in cybernetic terms, but psychological goals in particular are just one of the specific types of variable controlled by the brain (DeYoung & Weisberg, 2018). Unlike some psychologists who define psychological goals in terms of conscious representations (e.g., Elliot & Fryer, 2008), we assert that they are not necessarily conscious. However, their pursuit must involve processes over which the individual can, in principle, exert voluntary control, which entails that they must be pursuable via output through the voluntary muscular system or through the operation of selective attention and working memory. (Note, however, that pursuing psychological goals will also always involve involuntary processes as well, such as those involved in many aspects of perception.) Under this definition, the direct control of blood pressure by the brain and kidneys does not involve a psychological goal, but voluntary control of blood pressure as measured with a blood-pressure cuff, through modifications of diet, exercise, or medication, does. To provide a hypothetical example of an unconscious psychological goal, one might work extra hard to please a coworker who reminded one of one's father, even if the resemblance was not consciously recognized and even if one did not realize that one was especially motivated to please this coworker. In what follows, whenever we use the term *goal*, it is in reference to psychological goals.

Goals are hierarchically organized, in that complex goals are pursued by specifying simpler subgoals, which are themselves composed of even simpler subgoals, down to the level of specific motor or cognitive operations (Figure 1; Carver & Scheier, 1998). One's goal hierarchy is unlikely to be perfectly integrated under a small number of goals at the highest level because goals are often in competition or conflict with one another; pursuing one goal often interferes with the pursuit of others (Mansell, 2005). Psychological functioning, therefore, frequently requires compromises to satisfy as many of one's goals as possible without undermining any that are particularly important. Obviously, not all of one's goals are actively guiding behavior or decision making in any given situation, but all goals that are retained in memory (whether or not they are readily consciously accessible) may contribute to a person's ongoing interpretations of the world and are potentially relevant for function and dysfunction. Every person has stored in memory an extensive, though typically somewhat vague and conflicted, set of goals, which constitute a representation of the desired future and which are crucial for guiding behavior and interpreting experience. Identifying psychopathology requires an assessment of the degree to which people are successfully pursuing their collection of goals (For consideration of cases in which successful goal pursuit harms other people, see the *Agreeableness versus Antagonism* section).

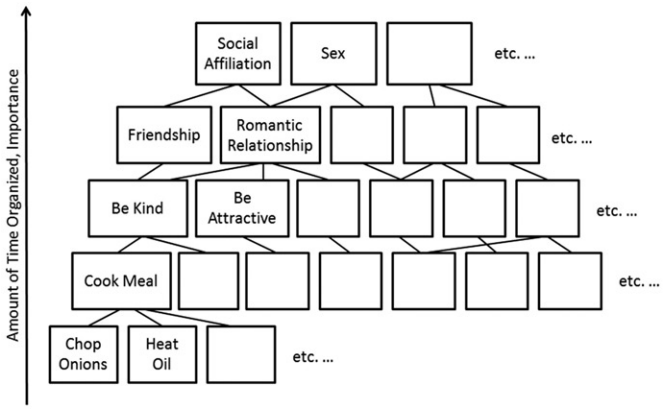

Figure 1. Hierarchical structure of goals for a hypothetical individual. *Note.* Lines between levels indicate dependence of goals on subgoals and indicate (a) that goals often require multiple subgoals and (b) that goals sometimes advance multiple superordinate goals. Goals higher in the hierarchy organize larger stretches of time and tend to be more important than those lower in the hierarchy. Goals within a row may also vary in importance or priority, however, and will tend to organize differing amounts of time accordingly. Severity of psychopathology increases with the number and importance of the goals that are disrupted. Adapted from On the Self-Regulation of Behavior by C. Carver and M. Scheier, 1998, Figure 5.2. New York: Cambridge University Press.

Any given goal organizes behavior over a certain period of time, and goals higher in the hierarchy, by definition, organize larger periods than goals linked beneath them in the hierarchy. The goal of writing an article, for example, may play a key role in organizing behavior over a period of weeks or months; the goal of getting a Ph.D. organizes behavior for 5 years or more. Asking a particular person out on a date is a relatively brief goal; the goal of establishing and maintaining a successful romantic relationship organizes a much longer period. Dysfunction can be assessed in terms of persistent failure to pursue one's goals successfully, with its severity determined by how much time the failed or endangered goals are organizing.

The organizational capacity of goals can be described in terms of psychological entropy (DeYoung, 2013; Hirsh, Mar, & Peterson, 2012). Entropy is a quantitative measure of disorder that was first formulated in relation to thermodynamics and mechanical systems but was later generalized to all information systems, for which it describes the uncertainty or unpredictability of the system (Shannon, 1948).[5] The entropy of a cybernetic system at any given time corresponds to uncertainty regarding its capacity to move toward its goals (Wiener, 1961). This is dependent on the uncertainty of the three cybernetic elements: goals, representations of the world, and operators. Are nonconflicting, achievable goals governing behavior? Is the world being interpreted correctly? Will the operators succeed in moving toward the goals? *Psychological entropy* reflects the number of plausible options or affordances available to the individual for interpretation (both perceptual and abstract) and for action (Hirsh et al., 2012). In other words, the harder it is to answer the questions "What is happening?" and "What should I do?" the higher the level of psychological entropy. These are not necessarily conscious questions, but rather assessments carried out by the brain unconsciously as well as consciously, through pervasive mechanisms of predictive coding (Clark, 2013; Friston, 2010).

[5]Note that the invocation of "disorder" suggests that "mental disorder" could reasonably be interpreted abstractly in terms of entropy. For the purposes of the present discussion, however, we choose to retain the more typical meaning of "mental disorder" as an official diagnostic entity.

Table 1 Definitions of key terms.

Cybernetics	The study of principles governing goal-directed systems that self-regulate via feedback.
Psychological goal	A representation (conscious or unconscious) of the desired state of some variable, capable of being pursued via output through the voluntary muscular system or through the operation of selective attention and working memory.
Psychopathology	Persistent failure to move toward one's psychological goals due to failure to generate effective new goals, interpretations, or strategies when existing ones prove unsuccessful.
Mental disorder	Psychopathology deemed sufficiently severe for diagnosis. (What constitutes sufficient severity is sociopolitically negotiated.)
Personality	All reasonably enduring psychological individual differences.
Personality traits	Probabilistic descriptions of relatively stable patterns of emotion, motivation, cognition, and behavior, in response to classes of stimuli that have been present in human environments over evolutionary time.
Characteristic adaptations	Relatively stable goals, interpretations, and strategies, specified in relation to an individual's particular life circumstances.

Psychopathology will typically be associated with high levels of psychological entropy, given that it is characterized by difficulty in pursuing goals effectively. For two reasons, however, it is not the case that any increase in psychological entropy is indicative of dysfunction or psychopathology. First, there is the second law of thermodynamics: Entropy naturally tends to increase over time in all systems. Living systems maintain relatively low entropy states by throughput of energy, which entails that, although they inevitably experience frequent gains in entropy, they are also able to reduce entropy again through work (Friston, 2010; Schrödinger, 1944/1967). Cybernetic dysfunction occurs only when the system persistently fails to be able to reduce entropy. To put this in more human terms, the world changes unpredictably, and people themselves change over the course of development, rendering new adaptations necessary; these changes lead to psychopathology only when the person fails to be able to adapt and cope with them.

Second, psychological entropy or uncertainty has the remarkable property of being both innately threatening and innately rewarding (DeYoung, 2013; Gray & McNaughton, 2000; Peterson & Flanders, 2002). The reason it is innately threatening is obvious; an organism's fitness depends on keeping entropy at a level that interferes minimally with fitness-relevant goals. The reason it is innately rewarding is less obvious but reflects precisely the fact that organisms inhabit complex, changing, and unpredictable environments. Given this situation, existing knowledge and strategies often prove inadequate, and new ones must be learned, which requires exploration. Rather than waiting to explore only when entropy increases spontaneously, it is advantageous to explore voluntarily, which means intentionally increasing the entropy of the system, with the expectation that one will be able to reduce it successfully again, having learned new adaptations (DeYoung, 2013, 2015; Schwartenbeck, FitzGerald, Dolan, & Friston, 2013). The fact that uncertainty is innately rewarding as well as threatening entails that people actively seek out situations of increased entropy. As examples, consider the appeal of gambling, novel aesthetic experiences, and challenging sports.

In sum, neither occasional setbacks in one's progress toward one's goals nor occasionally placing oneself in situations that increase uncertainty about whether one can achieve one's goals is sufficient to identify psychopathology. Only when the increased psychological entropy involved in these situations cannot be decreased again given the individual's existing set of goals, interpretations, and strategies, *and* the individual proves unable to generate new goals, interpretations, or strategies that allow resumption of successful goal pursuit, is psychopathology present. (We prefer the term *strategies* to *operators*, in the psychological context, for reasons that are clarified in the next section.) Hence we arrive at our definition: *Psychopathology is persistent failure to move toward one's goals due to failure to generate effective new goals, interpretations, or strategies when existing ones prove unsuccessful* (Definitions of key terms are listed in Table 1).

The Relation of Psychopathology to Personality

CB5T asserts that personality encompasses all reasonably enduring psychological individual differences (including those labeled "abnormal" or "pathological") and that all of those can be classified as traits or characteristic adaptations. This distinction between traits and characteristic adaptations is crucial for understanding psychopathology because our theory asserts that having extreme levels of a trait is not, in itself, sufficient for the identification of psychopathology. Although psychopathology typically is associated with extremes in one or more traits, it need not be, and failure of characteristic adaptations is necessary for the presence of psychopathology. Before elaborating this argument, we first review CB5T's definitions of personality traits and characteristic adaptations.

Characteristic Adaptations and Personality Traits

We begin with characteristic adaptations because they are closest to the cybernetic concepts discussed so far. According to CB5T, "characteristic adaptations are relatively stable goals, interpretations, and strategies, specified in relation to an individual's particular life circumstances" (DeYoung, 2015, p. 38). The first part of this definition incorporates the three necessary components of any cybernetic system, with *interpretations* being representations of the state of the world (in terms of potential motivational significance as well as fact) and *strategies* being organized collections of operators. The second part specifies that, to be characteristic adaptations, these components must result from the specific learning history of the individual. They are the updateable memory contents of the cybernetic system. Characteristic adaptations are not evolutionary adaptations (which reflect changes to the genome from generation to generation); rather, they are learned adaptations, acquired during an individual lifetime in response to experience. They reflect the characteristic ways that the individual has

adapted to his or her particular circumstances. To be "characteristic" in this context simply means to be typical of the person. This raises the point that not every adaptation one makes at a given moment is characteristic. A goal, strategy, or interpretation one adopts only briefly or in a single situation, never to adopt it again, is certainly an adaptation to that situation, but it is not a *characteristic* adaptation because an adaptation must be reasonably persistent to count as part of personality. It is when one's characteristic, enduring adaptations to the world prove to be maladaptations, and one is unable to replace them with better ones, that psychopathology emerges.

Whereas characteristic adaptations are constructs specified explicitly in relation to an individual's particular cultural or idiosyncratic circumstances, traits are more universal. According to CB5T, personality traits are probabilistic descriptions of relatively stable patterns of emotion, motivation, cognition, and behavior, in response to classes of stimuli that have been present in human environments over evolutionary time (DeYoung, 2015). That traits are probabilistic means that they describe the likelihood of being in the states associated with the trait in question. Someone scoring high in Neuroticism, for example, will not be in a constant state of negative emotion but will experience negative emotion more often and more intensely than someone low in Neuroticism. Scores on personality trait questionnaires correspond very well to the average of states measured repeatedly over time, despite considerable within-person variability in states over time (Finnigan & Vazire, 2018; Fleeson & Gallagher, 2009).

The within-person variability that renders traits probabilistic is not purely random; it reflects the features of the situations in which people find themselves. For each trait, there is a relevant class of such features (stimuli) to which human beings have evolved to respond (e.g., threat-related stimuli for Neuroticism). Trait-relevant states can be evoked by both external and internal features of the situation (such that the person high in Neuroticism may often experience threat simply due to ruminating on various concerns). This means that traits are not completely decontextualized, as some theorists have claimed, but rather reflect responses to contextual features. Nonetheless, the classes of stimuli to which trait mechanisms respond—such as threats, rewards, distractions, or other people—are typically so broad as to be present in many diverse situations (Funder, 1991).

Most importantly, these classes of stimuli have been present over evolutionary time, such that the brain's cybernetic mechanisms have evolved to respond to them effectively. Traits, therefore, reflect variation in the parameters of the cybernetic mechanisms that evolved to create and enact one's goals, interpretations, and strategies. In other words, the regularities in behavior described by traits are not the result of a goal or set of goals acquired from the individual's specific life experiences (those would be characteristic adaptations); rather, they are due to regularities in the functioning of the cybernetic mechanisms that generate and pursue those goals. Neuroticism, for example, does not reflect one's learned response to a particular stressor but rather reflects one's tendency toward defensive responses to any stressor.

Traits reflect varying parameters of universal human mechanisms, whereas characteristic adaptations are learned responses to specific circumstances. These two categories of psychological individual differences are causally linked in three ways. First, the mechanisms that underlie traits are necessary for carrying out characteristic adaptations in moment-to-moment behavior. Second, differences in traits lead to the development and adoption of different characteristics adaptations (e.g., the goals, interpretations, and strategies of an extraverted person are likely to be systematically different from those of an introverted person). Third, characteristic adaptations may create circumstances that cause changes in traits (e.g., taking a job in sales might eventually cause someone to become more extraverted). Although traits are remarkably stable, they can and do change over time (Roberts, Wood, & Caspi, 2008). Even in adulthood, systematic changes of traits in response to specific environmental changes have been documented (Bleidorn, Hopwood, & Lucas, 2018).

One potential source of confusion in CB5T is that not all psychological goals are characteristic adaptations (DeYoung & Weisberg, 2018). The fact that traits encompass patterns of motivation (in addition to emotion, cognition, and behavior) suggests that some goals are more closely linked to traits than to characteristic adaptations. In particular, such goals must be innate (e.g., goals of avoiding punishment or affiliating with other people) rather than learned (although learning processes may adjust their motivational intensity), such that they reflect cybernetic mechanisms present in any intact human brain. Note that their presence as goals in the evolved human brain does not mean they must be equally important to each individual. Indeed, it is precisely their variation in importance and intensity in the population, due to both genetic and environmental forces, that qualifies them as personality traits. Their level in the individual is what constitutes the trait, rather than their mere existence in the individual (whereas some characteristic adaptations may be defined in a binary manner in terms of their existence; e.g., I am either a lawyer or I am not). Nonetheless, behavior that can achieve these very broad goals always involves characteristic adaptations as well. One does not go about pursuing the goal of affiliating with another human being, for example, without specifying various more fine-grained subgoals and strategies that reflect the specifics of one's situation. Thus, even if we can identify that a person is failing in the pursuit of some very broad species-typical goal (what we might call a basic human need), that failure will also be reflected in the failure of the characteristic adaptations necessary to pursue that broad goal. Hence, we retain our criterion that psychopathology involves failure of characteristic adaptations.

We close this section with some examples to help clarify the distinction between traits and characteristic adaptations:

> Being argumentative is a trait; being a trial lawyer is a characteristic adaptation. Liking to frolic with friends is a trait; belonging to a fraternity is a characteristic adaptation. Being

> prevention focused is a trait; checking the stove every time one leaves the house is a characteristic adaptation. Having an avoidant coping style in general is a trait; habitually avoiding a particular acquaintance is a characteristic adaptation. (DeYoung, 2015, p. 40)

Note that characteristic adaptations, such as those listed here, often consist of collections of related goals, interpretations, and strategies, but they are always decomposable into those three elements.

Maladaptation and the Identification of Psychopathology

To the extent that characteristic adaptations do not allow people to pursue their goals successfully, yet are not abandoned or replaced with more effective adaptations, they may be considered *maladaptations*. Maladaptive behaviors can persist for a variety of reasons, including that they are so well-learned as to be thoroughly entrenched habits; that fear or anxiety prevents the exploration that would be necessary to develop new adaptations; that one is too easily daunted or distracted to pursue a challenging goal successfully, despite knowing viable strategies; that the goals one has adopted are fundamentally incompatible with one another or with one's basic needs; or that one is unable to coordinate one's goals with other people's goals effectively (DeYoung, 2015; Mansell, 2005). Per the definition in Table 1, the persistence, for whatever reason, of unsuccessful characteristic adaptations—goals, interpretations, and strategies—is central to psychopathology.

Our specification that psychopathology involves failure of characteristic adaptations is reasonably congruent with recent editions of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) of the American Psychiatric Association, which uses impairment in social or occupational functioning as a criterion for disorder. (Note, however, that the *DSM* definition of mental disorder does not provide specific definitions of "impairment," "disability," or "dysfunction"; Stein et al., 2010.) Merely having extreme levels of some trait is not sufficient to indicate psychopathology; one must also be failing to achieve one's goals. This is true even for traits identifiable as "persistent symptoms" in a clinical context, unless the symptom in question explicitly describes dysfunction (some failure of goal pursuit). Hence, carrying out an assessment of psychopathology must always involve an attempt to determine the range of the person's goals, situated in his or her particular life context (cf. Mansell, 2005; Widiger & Trull, 2007). This will include attempting to determine the relative importance placed on various goals over time.

We take the specifics of a person's goal hierarchy to be a matter of fact about which there is an objective answer (which can change over time), despite the fact that the person cannot have perfect subjective (conscious) awareness of all specifics of his or her own goals. Ultimately, these goals are representations of a desired future state that are physically instantiated within the brain. The fact that neither the clinician nor the person being evaluated will have full direct access to the goal hierarchy in question will often introduce some degree of uncertainty into the process of evaluating psychopathology, but that does not contradict our claim that its presence is fundamentally an objective matter of fact rather than determined by socially negotiated value judgments. What matters is whether people are successfully pursuing their own goals, even those they have not consciously selected and may not be aware of. Of course, the individual's social context will have considerable effect on what goals have been adopted (and may have encouraged the adoption of unrealistic or conflicting goals), but that is separate from the question of whether the individual is functioning well cybernetically—that is, capably pursuing whatever goals he or she happens to have adopted. Society, therefore, can contribute causally to psychopathology, but it does not provide criteria for psychopathology.

The fact that our theory specifies that dysfunction is to be identified by criteria within the individual renders dysfunction largely overlapping with harm. Although Wakefield (2007) defined "harmful" as "judged negative by sociocultural standards" (p. 149), more common definitions emphasize that what is harmful causes physical or mental injury to the person. In cybernetic terms, harm and injury necessarily involve interference with the goals of the organism (psychological or otherwise). The severity of dysfunction, therefore, should correspond reasonably well to the severity of harm. This entails that, although harm is often subjectively obvious, people can be harmed without awareness.

Because harm is often subjectively felt as negative emotion, something important to consider in relation to our contention that extreme levels of a trait do not, by themselves, indicate psychopathology is that most people have an explicit goal of not being miserable. This entails that very high levels of Neuroticism (or any of its various facets, such as anxiety or depression), which describes the tendency to experience negative emotion, are almost always associated with some level of psychopathology. Nonetheless, it is possible for people who are highly neurotic to have accepted their tendency toward negative emotion and to have developed characteristic adaptations that allow them to pursue their goals effectively, despite frequent experiences of negative emotion. To the degree that they have abandoned the goal of avoiding negative emotion, they may not be dysfunctional. This state is presumably achieved by few, and most people high in Neuroticism are likely to have some degree of psychopathology, though it often will not be severe enough to be deemed mental disorder.

Extreme trait levels are neither necessary nor sufficient for psychopathology, but they are often associated with psychopathology because they can lead to persistent failure of characteristic adaptations and cybernetic dysfunction. Because traits represent functional parameters of the mechanisms that allow us to generate and enact our characteristic adaptations, extreme values of those parameters can lead to instability in the system, increased psychological entropy, failure of immediate goal pursuit, and inability to generate effective new adaptations, all of which risk serious disruption of goal pursuit. Feedback processes are also possible, as failure of characteristic adaptations can lead to changes in

traits, most obviously increases in Neuroticism. A major goal of our theory is to describe the underlying mechanisms by which particular traits are linked to dysfunction, but first we must review how the surface features of personality and psychopathology are empirically related.

Hierarchical Covariance Structure in Personality and Psychopathology

Having explained the theoretical relation of psychopathology to personality in general, we now switch our focus to the empirical relation between specific dimensions of psychopathology and personality traits. This requires an understanding of what is meant by "hierarchical covariance structure." We have already encountered a different application of the concept of hierarchy; goals are organized hierarchically within the individual in two senses: First is a hierarchy of timescale, in which broader, longer term goals must be accomplished through narrower, shorter term subgoals (Figure 1). Second, goals are prioritized for each individual, even when they are on a similar timescale, such that some are more important, and therefore more likely to influence thought and behavior (and hence to organize more of the person's time), than others. Another type of hierarchy is crucial for understanding personality and psychopathology, however, and this is the hierarchical structure of patterns of covariance among traits (including persistent or recurring symptoms of psychopathology)—in other words, which traits tend to appear in the same individuals. At each level of the hierarchy, traits are grouped together because they tend to covary with each other more strongly than with the other traits represented at the same level (e.g., Neuroticism groups together anxiety, depressed mood, irritability, and emotional lability). At the next level down in the hierarchy, however, some of these traits are separated, again based on which are most closely related (e.g., below Neuroticism, anxiety and depressed mood can be grouped in one dimension, whereas irritability and emotional lability can be grouped in another; see the *Neuroticism, Negative Affect, and Internalizing* section). At each level of the hierarchy below the highest, each dimension contains both valid variance that is shared with other traits at the same level (allowing identification of higher dimensions) and valid unique variance that is not shared, and this is true both phenotypically and genetically (Jang, McCrae, Angleitner, Riemann, & Livesley, 1998; Jang et al., 2002; McCrae et al., 2008).

Before characterizing the hierarchical covariance structure of the dimensions of psychological individual differences, it is important to establish that psychopathology is dimensional rather than categorical. Mental disorders are traditionally and officially considered to be discrete binary categories, discontinuous from healthy personality variation. However, no mental disorder has ever been empirically verified as a categorical entity (Carragher et al., 2014; Haslam, Holland, & Kuppens, 2012; Markon & Krueger, 2005; Walton, Ormel, & Krueger, 2011; Widiger & Samuel, 2005; Wright et al., 2013). Treating assessments of mental disorder as dimensional (by using symptom severity scores rather than binary diagnoses) increases their reliability and validity substantially (Markon, Chmielewski, & Miller, 2011). Finally, and most directly relevant to the link between psychopathology and personality, assessments of psychopathology have been shown to be measuring the same underlying latent trait dimensions as normal personality measures, and this is true not only of the so-called personality disorders but of mental disorders more generally (Griffith et al., 2010; Markon, Krueger, & Watson, 2005; Suzuki, Samuel, Pahlen, & Krueger, 2015; Stepp et al., 2012; Widiger & Trull, 2007).

This last observation is what is crucial for the development of a theory of psychopathology based on CB5T, which is a theory of the mechanisms underlying the Big Five. The Big Five personality dimensions have been demonstrated extensively in both self- and other-ratings of descriptors drawn either from the natural language (trait-descriptive adjectives) or from existing questionnaires not specifically designed to measure the Big Five (John et al., 2008; Markon et al., 2005; Waller, DeYoung, & Bouchard, 2016). They thus form a solid basis for a theory of personality. Five similar dimensions have been repeatedly observed in the patterns of covariation of symptoms and diagnoses of mental disorders, which have been labeled *Negative Affect* or *Internalizing*, *Disinhibition*, *Antagonism*, *Detachment*, and *Psychoticism* or *Thought Disorder* (Kotov et al., 2017; Krueger & Markon, 2014; Wright & Simms, 2015; cf. Harkness, Reynolds, & Lilienfeld, 2014). Indeed, as previously mentioned, psychometric evidence indicates that these are actually the same latent dimensions—at least for four of the Big Five. The one dimension that does not map as cleanly onto one of the Big Five is Psychoticism. However, as we discuss further later, Psychoticism can be mapped cleanly onto the Big Five hierarchy at the level below the Big Five. After splitting the Openness/Intellect dimension into subdimensions of Openness to Experience and Intellect, Psychoticism is seen to be a maladaptive variant or facet of Openness specifically, unrelated or even negatively related to Intellect (DeYoung, Carey, Krueger, & Ross, 2016; DeYoung, Grazioplene, & Peterson, 2012).

CB5T makes use of a four-level trait hierarchy (Figure 2). The two most familiar of these levels are probably the Big Five and their many facets at the bottom level. No consensus exists regarding the number of facets within each Big Five dimension, but empirical evidence indicates that exactly two factors (labeled *aspects*) exist between each of the Big Five and their facets (DeYoung, Quilty, & Peterson, 2007; Jang et al., 2002). For the purposes of the current theory we focus primarily on the Big Five, as the best established level, and secondarily on the aspect and metatrait levels. It may be possible to insert additional levels of the hierarchy between the Big Five and the metatraits or below the facets, but we focus on those labeled in Figure 2, for which there is good empirical evidence in the study of normal personality (DeYoung, 2006; DeYoung et al., 2007; McCrae et al., 2008).

One important caveat regarding the schematic representation of this hierarchy in Figure 2 is that it may falsely imply

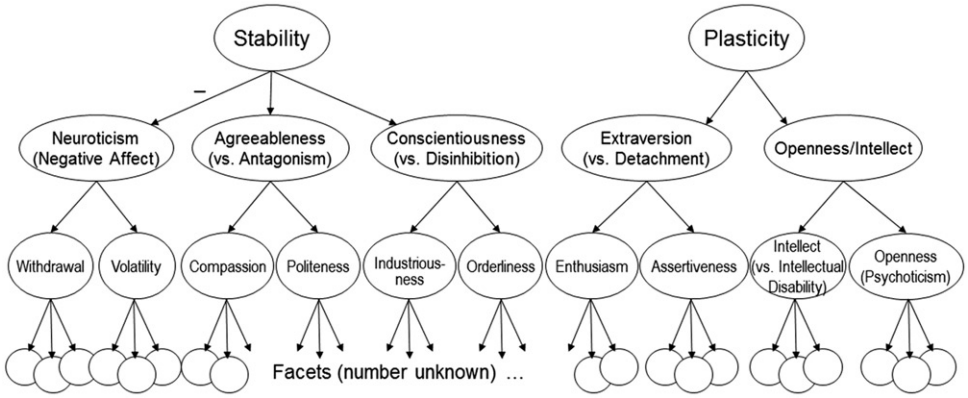

Figure 2. A hierarchical taxonomy of personality traits, with the five major dimensions of psychopathology (plus intellectual disability) indicated in parentheses. *Note.* Top level: metatraits. Second level: Big Five. Third level: aspects. Bottom level: facets. (Constructs at every level are traits.) Negative sign indicates inverse association of Neuroticism with Stability.

simple structure that does not exist, especially below the level of the Big Five. The lines connecting traits in the figure suggest that any traits subsumed under Plasticity should be unrelated to any traits subsumed under Stability, but this is not the case. For example, Enthusiasm is positively related to Compassion, and Assertiveness is negatively related to Politeness. It has long been known that personality does not have simple structure, and cross-relations not depicted in the figure are likely to be meaningful (e.g., the covariation between Assertiveness and Politeness may, in part, reflect the influence of testosterone on personality; DeYoung, Weisberg, Quilty, & Peterson, 2013). The figure is a simplified heuristic.

Cybernetic Mechanisms Underlying Dimensions of Psychopathology

Given that our definition of psychopathology specifies failure of characteristic adaptations as a criterion and that extreme levels of personality traits are not necessary for psychopathology, one may wonder why manifestations of psychopathology can be described so well in terms of personality traits. In part, this is because characteristic adaptations are specific to individuals or groups of individuals, whereas, when attempts are made to specify symptoms for diagnosis, a usable system must be generally applicable to all people to whom the diagnosis might be applied. Thus, it is precisely the forms of psychological dysfunction that might potentially apply to anyone that get codified in diagnostic systems, and these are, by our definition, traits, as long as they are reasonably persistent or recurring.

More substantively, we take the correspondence between psychopathology and personality to reflect that traits are caused by the typical functional levels of the mechanisms that are necessary to generate and carry out one's characteristic adaptations. These are the cybernetic mechanisms that govern moment-to-moment behavior. Whatever goal one is pursuing, the same set of core mechanisms are likely to be involved. When persistent cybernetic dysfunction occurs, it is typically because one (or more) of these mechanisms is not functioning well as an integrated component of the system, and this malfunctioning often stems from an extreme value of some parameter of the mechanism. An extreme trait, therefore, typically presents a persistent risk for dysfunction.

Cybernetic control of behavior can be described heuristically in terms of a cycle, and the stages of this cycle usefully delineate the core tasks that the human cybernetic system must accomplish (DeYoung, 2015): (a) *Goal activation*: A goal must be sufficiently activated to control behavior; (b) *Action selection*: An operator deemed likely to move the system toward the goal must be chosen from the available repertoire; (c) *Action*: The operator must be enacted; (d) *Outcome interpretation*: The current state of the world is assessed; (e) *Goal comparison*: The current state is compared to the goal state. If they match, then the goal is accomplished and a new goal will be activated. If a mismatch is detected, however, the cycle begins again with the same goal, and a new action must be selected (or alternatively the goal may be abandoned). In an important sense, this linear cycle is merely a convenient heuristic; it suggests that these operations are carried out serially, when, in reality, many of the processes associated with most stages of the cycle take place in parallel more or less constantly (DeYoung & Weisberg, 2018). For example, we are almost constantly interpreting the world around us, not just at the moment of completing some action, and thus mismatches between our current state and our desires or expectations can be detected at any time. Nonetheless, the stages of the cycle are helpful for conceptualizing the core mechanisms that constitute the human cybernetic system.

In the rest of this section, we describe which mechanistic parameters are likely to underlie each trait dimension and how and why their extreme values are linked to psychopathology, considering neurobiological as well as psychological evidence. Throughout this section, we draw on a study showing that the 25 scales of the Personality Inventory for *DSM-5* (PID-5) share the same 10-factor space as the Big Five Aspect Scales, which were designed to assess the 10 empirically identified aspects depicted in Figure 2 (DeYoung et al., 2016; Krueger, Derringer, Markon, Watson, & Skodol, 2012). Although these 25 symptom dimensions, or "pathological personality traits," were designed specifically to assess personality disorders in the *DSM*, they describe many

or even most of the symptoms that are present in other mental disorders as well (Wright & Simms, 2015). Thus, they are informative for linking the Big Five hierarchy to psychopathology in general. Linking the PID-5 scales to the aspect-level traits is additionally informative beyond the well demonstrated links between the PID-5 and Big Five (e.g., Gore & Widiger, 2013; Krueger & Markon, 2014).

We also draw on a recent review of the neurobiological correlates of the traits appearing in Figure 2, which framed them in terms of CB5T (Allen & DeYoung, 2017). Although CB5T attempts to identify a reasonably coherent cybernetic function associated with each dimension, it acknowledges that each of those functions will be carried out by a complex set of neural mechanisms. This complexity is implied not only by a network perspective on brain function (Yeo et al., 2011) but also by evidence that most behavioral traits are massively polygenic, influenced by variation in hundreds or even thousands of genes (Manolio et al., 2009; Munafò & Flint, 2011).

Neuroticism, Negative Affect, and Internalizing

Neuroticism reflects the general tendency to experience negative emotions of all kinds, with accompanying cognitive processes like rumination and self-doubt. Neuroticism is difficult to distinguish statistically from the general risk factor for the *internalizing* disorders (Griffith et al., 2010), one of the two major dimensions of covariation among common mental disorders (the other being *externalizing*; Krueger et al., 2007). Internalizing encompasses disorders of anxiety, depressive mood, and fear (Kotov et al., 2017). CB5T, like several other theoretical accounts of Neuroticism, specifies that this dimension reflects sensitivity of defensive responses to threat, punishment, and uncertainty (DeYoung, 2015; Shackman et al., 2016). In cybernetic terms, a threat is any indication that progress toward a goal may be hindered or prevented, and punishment is any definite thwarting or failure of a goal (and most punishments also serve as threats of further punishment). As just discussed, uncertainty is innately threatening because it increases psychological entropy (Gray & McNaughton, 2000; Hirsh et al., 2012). Indeed, all threats inherently increase psychological entropy. In terms of the cybernetic cycle, Neuroticism is most relevant to the final stage, at which it corresponds to how likely a mismatch is to be detected. The highly neurotic person chronically feels that things are not as they should be, that the current state does not match the desired state. Neuroticism will influence other stages of the cycle as well, however, including which goals are activated (e.g., a higher proportion of avoidance goals) and how information is weighted when selecting an action.

Defensive responses initiated after threat detection are of two distinct kinds, which can be described as active defense and passive avoidance (DeYoung, 2015; Gray & McNaughton, 2000). CB5T posits that the two aspects of Neuroticism—Volatility and Withdrawal—correspond to these two forms of defensive response, respectively. Active defense involves emotional and behavioral responses to immediate threats where the only motivation is to escape or eliminate the threat. Volatility encompasses emotional lability, irritability, and anger (DeYoung et al., 2007; 2016). Passive avoidance, in contrast, involves involuntary inhibition of approach toward a goal in response to increases in psychological entropy. Thus, it occurs when motivation is conflicted, most commonly in approach-avoidance conflicts, in which an approach goal (e.g., acquiring a romantic partner) conflicts with an avoidance goal (e.g., avoiding rejection), creating uncertainty about the action to select (Gray & McNaughton, 2000; Mansell, 2005). Passive avoidance can be subdivided into anxiety and depression, which cover most of the Withdrawal aspect of Neuroticism. (Here "depression" refers specifically to depressed mood, negative self-evaluation, and hopelessness, not to the broader criteria for an official diagnosis of a mood disorder.) The label "Withdrawal" does not refer to "social withdrawal" (which is related more specifically to low Extraversion or Detachment) but to the involuntary withdrawal of effort from a goal, which constitutes passive avoidance (DeYoung, 2015).

Anxiety describes the initial response to goal conflict and uncertainty, in which the perceived possibility of reward has not been completely overcome by the likelihood of punishment, and thus the goal in question is still perceived to be potentially attainable. In addition to inhibition or slowing of approach to the conflicted goal, anxious passive avoidance involves increased attention to both sensory input and information in memory in order to scan for further threat (Gray & McNaughton, 2000; Hirsh et al., 2012). In addition, during anxiety, arousal increases to prepare to switch to active defense if danger becomes too great. These defensive measures help to prevent encounters with danger that might be associated with the goal in question. In depression, the goal is perceived to be unattainable and approach motivation is extinguished (Carver & Scheier, 1998). This should lead to the abandonment of an unreachable goal and the selection of a new goal, but it can overgeneralize, leading to the dysfunctional extinction of many goals characteristic of clinical depression. The extremely high comorbidity of anxiety and depression is probably due to the facts that both are forms of passive avoidance and that people often oscillate between feeling that a conflicted goal is still potentially obtainable versus out of reach.

The distinction between Volatility and Withdrawal appears to be important for distinguishing different types of psychopathology. Within a sample of 275 patients diagnosed with mood disorders, Volatility specifically predicted a diagnosis of bipolar disorder, whereas Withdrawal specifically predicted a diagnosis of unipolar disorder (Quilty, Pelletier, DeYoung, & Bagby, 2013; cf. Stanton, Gruber, & Watson, 2017). Given a general risk for internalizing psychopathology conferred by high Neuroticism, the balance of its two aspects may influence whether unipolar or bipolar psychopathology develops. In addition, the tendencies toward anger and lability associated with Volatility seem likely to be associated not only with internalizing but also with externalizing problems, such as those described by intermittent explosive disorder or antisocial personality disorder. Considering these

traits from a developmental perspective, a similar distinction has been made in research on personality and mental health in childhood, between "anxious distress" and "irritable distress," which may be informative about risk for later disorder (Caspi & Shiner, 2006).

Brain structures that have been empirically linked to Neuroticism, with reasonably solid evidence, include the amygdala and adjacent bed nucleus of the stria terminalis, the hypothalamus (as part of the hypothalamic-pituitary-adrenal axis), and the rostral anterior cingulate cortex (ACC) and adjacent medial prefrontal cortex (Allen & DeYoung, 2017; Holmes et al., 2012; Shackman et al., 2016). These structures have also been implicated in internalizing disorders (Holmes et al., 2012; Shackman et al., 2016). Gray and McNaughton (2000) referred to the brain systems that govern active defense and passive avoidance as the fight–flight–freeze system and the behavioral inhibition system (BIS), respectively. The BIS is centered around the hippocampus and extended amygdala, whereas the fight–flight–freeze system is centered around the hypothalamus and brainstem (although it also involves the amygdala). These systems are both modulated by serotonin and norepinephrine, neurotransmitters that have been linked empirically to Neuroticism (Allen & DeYoung, 2017; Gray & McNaughton, 2000). The serotonergic system is the most common target of pharmacological treatments for anxiety and depression, and reductions in the symptoms of these disorders following treatment with selective serotonin reuptake inhibitors (SSRI) are mediated by declines in Neuroticism (Du, Bakish, Ravindran, & Hrdina, 2002; Quilty, Meusel, & Bagby, 2008; Tang et al., 2009). Of interest, some evidence suggests that SSRIs may be more effective for reducing characteristics related to Volatility than to Withdrawal (Ilieva, 2015; Kamarck et al., 2009).[6]

More than any of the other four dimensions, elevated Neuroticism is likely to be a result as well as a cause of psychopathology. Cybernetic dysfunction is likely to result from too frequent engagement in passive avoidance or active defense, but, additionally, distress in response to increased psychological entropy due to failure in goal pursuit is the final common pathway of nearly all psychopathology. Few traits posing a risk for psychopathology are associated with low Neuroticism—that is, with insensitivity of threat detection and lack of distress. The fearless and unemotional traits linked to psychopathy may be one exception (Henry, Pingault, Boivin, Rijsdijk, & Viding, 2016; Miller & Lynam, 2003; Patrick, Fowles, & Krueger, 2009). Also, without causing psychopathology per se, very low Neuroticism might be associated with dangerous levels of risk taking.

[6]Although the serotonergic system is the most common target of pharmacological treatments for anxiety and depression, which are the major components of Withdrawal, we note that its causal action in treatment is poorly understood, so much so that debate remains regarding whether serotonergic function is typically reduced or elevated in most cases of depression (Andrews, Bharwani, Lee, Fox, & Thomson, 2015). Acute serotonin administration often increases anxiety in animal models, and one study found that a week-long course of a serotonin reuptake inhibitor (SSRI) increased, rather than decreased, neural reactivity to emotionally negative stimuli in people high in Neuroticism (Di Simplicio, Norbury, Reinecke, & Harmer, 2014).

Extraversion versus Detachment

Whereas Neuroticism reflects sensitivity to threat and punishment, Extraversion reflects sensitivity to reward (DeYoung, 2015; Wacker & Smillie, 2015). Extraversion encompasses a range of traits including talkativeness, assertiveness, sociability, positive emotionality, activity level, and excitement seeking. It is often manifested in social behavior because many human rewards are social in nature, involving innately rewarding classes of stimuli such as social affiliation, social status, and sex. However, the sensitivity of extraverts to reward is not limited to the social, as they respond more strongly to monetary and other rewards as well. Extraversion predicts better learning under conditions of reward in reinforcement learning paradigms, as well as facilitation of reaction times and accuracy following reward (Pickering, 2004; Robinson, Moeller, & Ode, 2010; Smillie, 2008).

In cybernetic terms, rewards are any stimuli that indicate progress toward or attainment of a goal, which renders the range of what people may find rewarding extremely broad (DeYoung, 2013, 2015). Rewards have both incentive properties (serving as cues for the possible attainment of a goal) and hedonic properties (enjoyment of goals when they are attained; Berridge, Robinson, & Aldridge, 2009). Because goals are nested, with larger goals achieved through subgoals, the achievement of a goal can be simultaneously a hedonic reward and an incentive indicating progress toward additional goals. Extraversion reflects variation in sensitivity to both incentive and hedonic rewards, and the difference between these two types may be reflected in the two aspects of Extraversion. Assertiveness seems more purely incentive related, reflecting the drive to achieve one's goals, and its link to social status is consistent with the fact that social status, or "dominance" in ecological terms, represents relative access to resources (DeYoung et al., 2013). Enthusiasm encompasses sociability and positive emotionality and is more related to the hedonic qualities of reward, though especially in an incentively rewarding context (DeYoung, 2015; Smillie, Geaney, Wilt, Cooper, & Revelle, 2013).

Considerable neuroscientific evidence links Extraversion to variation in the brain's reward system (Allen & DeYoung, 2017; Depue & Collins, 1999; DeYoung, 2013; Wacker & Smillie, 2015). At the heart of this system is the neurotransmitter dopamine, governing incentive reward. Dopamine is released in response to outcomes that are better than expected and triggers both approach behavior and learning about stimuli that predict reward. Gray (1973; Pickering & Gray, 1999) described the incentive reward system as the Behavioral Approach System (BAS). At least nine studies have used pharmacological manipulation of dopamine to show that Extraversion predicts variation in dopaminergic function (Allen & DeYoung, 2017). Studies of the brain's electrical activity also provide evidence that Extraversion is associated with the strength of the value signal encoded by dopamine (Wacker & Smillie, 2015). Both structural and functional magnetic resonance imaging (MRI) studies have linked Extraversion to key brain structures in the reward system, including the midbrain regions where dopaminergic

neurons originate, the nucleus accumbens and caudate nucleus, and the medial orbitofrontal cortex (Allen & DeYoung, 2017). The hedonic component of reward appears to be controlled by the endogenous opiates rather than by dopamine (Berridge et al., 2009), and one study has shown that Social Closeness (an indicator of Enthusiasm) moderates the emotional and behavioral effects of an opiate antagonist (Depue & Morrone-Strupinsky, 2005).

Psychopathology related to low Extraversion can be characterized as Detachment—inability to be motivated by the possibility of achieving goals or to enjoy rewards that are attained. Problems associated with Detachment involve anhedonia, amotivation or avolition, restricted affect, and lack of social connections (DeYoung et al., 2016; Kotov et al., 2017). Presumably because activation of the BIS directly inhibits the BAS (Gray & McNaughton, 2000), Extraversion is negatively correlated with Withdrawal, anxiety, and depression (DeYoung et al., 2007; Watson, Stasik, Ellickson-Larew, & Stanton, 2015). The opposite direction of causality is also possible, however, as being less sensitive to reward may be a particular risk for depression because lack of motivation to pursue goals can lead to hopelessness (Bress, Smith, Foti, Klein, & Hajcak, 2012; Pizzagalli et al., 2009; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008). Fewer forms of psychopathology are associated with high Extraversion—that is, with oversensitivity to reward—the most notable being mania (DeYoung, 2013; Stanton et al., 2017; Tackett, Quilty, Sellbom, Rector, & Bagby, 2008), but also the histrionic tendency toward attention seeking (Kotov et al., 2017; Samuel & Widiger, 2008; Wright & Simms, 2015). In addition, high Assertiveness specifically may be a risk for some externalizing problems (Watson et al., 2015).

Extraversion is particularly relevant to the first stage of the cybernetic cycle, where goals must become sufficiently activated to guide the system. People low in Extraversion may have difficulty finding many possibilities sufficiently energizing to trigger goal pursuit. Enthusiasm, in particular, with its emphasis on enjoyment of reward, is also relevant to the final stage of the cycle, at which the current state and goal state are compared. Those high in Enthusiasm are more likely to experience pleasure when things are going well. Like high Neuroticism, low Enthusiasm is probably more likely than most other traits to be caused by psychopathology because failure of goal-pursuit reduces opportunities for pleasure.

Conscientiousness versus Disinhibition

Conscientiousness reflects the tendency to be self-disciplined and organized. In developmental research, it is often described as "effortful control" (Caspi & Shiner, 2006). The two aspects of Conscientiousness are Industriousness, which involves the ability and tendency to work hard and avoid distractions, and Orderliness, which involves the ability and tendency to be neat, punctual, and concerned with details and rules. Both of these traits are crucial in the process of determining which goals should be followed and which actions selected at any given time. According to CB5T, the cybernetic function associated with Industriousness is to prioritize nonimmediate goals, whereas that associated with Orderliness is to follow rules set either by oneself or by others (DeYoung, 2015). Both of these functions require selecting and following through with strategies that are not merely responsive to the immediately rewarding or threatening properties of the situation but rather take more elaborate and abstract goals into account. Consistent with this observation, Conscientiousness predicts the ability to perform well on tasks that require decisions about how to prioritize various subtasks (Stock & Beste, 2015).

In relation to psychopathology, low Conscientiousness is most specifically associated with risk for disinhibited externalizing problems like impulsivity, substance abuse, criminal behavior, and attention deficit/hyperactivity disorder (Kotov, Gamez, Schmidt, & Watson, 2010; Kotov et al., 2017; Krueger et al., 2007). These problems appear to be primarily associated with low Industriousness, suggesting a tendency to prioritize immediate goals that interfere with longer-term goals (DeYoung et al., 2016). In contrast, Orderliness is associated positively with problems linked to compulsivity, such as rigid perfectionism and obsessive-compulsive personality disorder, suggesting an inflexible attention to detail and overreliance on rules (DeYoung et al., 2016; Samuel & Widiger, 2008). It thus appears that the risks associated with high versus low Conscientiousness are differentially associated with its two aspects, which may explain why indicators of disinhibition and compulsivity sometimes form distinct factors in analyses of the covariance structure of psychopathology (e.g., Forbes et al., 2017). Low Conscientiousness is a fairly general risk factor for psychopathology (Kotov et al., 2010), which is consistent with our definition of psychopathology, given that prioritizing goals and sequencing actions correctly is crucial for the accomplishment of all but the simplest goals.

Relative to Neuroticism and Extraversion, evidence for the neural substrates of Conscientiousness is sparse. Most clearly implicated both theoretically and empirically is the lateral prefrontal cortex. A number of structural MRI studies have reported positive associations between Conscientiousness and volume of regions in dorsolateral prefrontal cortex (DLPFC; Allen & DeYoung, 2017; Riccelli et al., 2017), which is in keeping with this region's role in forming and maintaining complex goals and plans (Bunge & Zelazo, 2006; Miller & Cohen, 2001). Conscientiousness is not associated with working memory ability, however, which is the function most commonly attributed to DLPFC and to the large cognitive control network of which DLPFC is a key node (DeYoung et al., 2005, 2009). However, there is another major brain network with a node in DLPFC, which encompasses networks typically labeled the *ventral attention* and *salience* networks (Yeo et al., 2011). This network appears to be crucial for prioritizing goals and directing attention away from distractions and toward goal-relevant stimuli, and recently found support for the hypothesis that it is an important substrate of Conscientiousness (Rueter, Abram, MacDonald, Rustichini, & DeYoung, 2018).

Other regions in this network that have been linked to Conscientiousness and Disinhibition include the insula and ACC (Abram et al., 2015; Allen & DeYoung, 2017). The lateral orbitofrontal cortex is not part of this network but has also been repeatedly linked to Conscientiousness and appears to be crucial for encoding the contrasting values of different outcomes, which is obviously relevant to prioritizing goals (Cheetham et al., 2017; Jackson, Balota, & Head, 2011; Matsuo et al., 2009; Nouchi et al., 2016; Rudebeck & Murray, 2011). Thus, a relatively coherent functional account is beginning to link Conscientiousness to neural substrates.

Agreeableness versus Antagonism

Agreeableness encompasses all of the traits directly related to altruism and cooperation. From a cybernetic perspective, Agreeableness has a unique role because it does not reflect variation in mechanisms that are necessary for pursuing every goal. Instead, it reflects variation in the mechanisms that are responsible for allowing people to coordinate their goals with those of others (DeYoung, 2015). Human beings are intensely social animals, so this capacity for coordination is fundamental to our nature and pervasively important in our development and adult functioning. Nonetheless, at least in adulthood, one can potentially manage to accomplish some of one's own goals with a minimum of attention to the needs and desires of others, and the most adaptive degree of Agreeableness is likely to vary with context. The two aspects of Agreeableness are Compassion, which reflects empathy, interpersonal concern, and the absence of callousness, and Politeness, which reflects the tendency to follow social norms and to avoid exploiting or acting belligerently toward others (DeYoung et al., 2007, 2016).

Psychopathology related to Agreeableness is mostly associated with a lack thereof, typically labeled "Antagonism" but also well characterized as selfishness. Antagonism is associated with various externalizing problems involving misconduct, antisocial behavior, and aggression, as well as with narcissistic entitlement and grandiosity, paranoid distrust, dishonesty, and inability to maintain stable social relationships (DeYoung et al., 2016; Kotov et al., 2017; Krueger et al., 2007; Wright & Simms, 2015). The initial presentation of CB5T argued that extremely low Agreeableness might indicate psychopathology even in the absence of cybernetic dysfunction, given that extremely disagreeable people are likely to cause suffering for others even if not for themselves (DeYoung, 2015). In further developing our theory of psychopathology based on CB5T, however, we have come to the conclusion that this exception is untenable. Psychopathology must involve failure to move toward one's own goals. Almost always, enough of one's own goals are sufficiently dependent on other people that dramatically mistreating people will tend to undermine those goals, leading to psychopathology. Nonetheless, we recognize the possibility of what has been called "the successful psychopath," someone who is competent but also sufficiently callous as to have little to no goal of affiliating with other people or preventing their suffering (Hall & Benning, 2006). This person may be able to attain his or her own goals while exploiting and hurting others. Such behavior is not inherently psychopathological, and preventing it is best considered a legal matter.

As with Conscientiousness, knowledge of the neural basis of Agreeableness is not yet extensive. It seems clear that variation in the systems responsible for social information processing, including emotional empathy and theory of mind (the ability to reason about the mental states of others), are involved, as one cannot accommodate the goals of others if one cannot understand them. Indeed, research has shown that some facets of Agreeableness predict performance on a theory of mind test (Allen, Rueter, Abram, Brown, & DeYoung, 2017; Nettle & Liddle, 2008). Theory of mind depends on components of the so-called default network, a set of brain regions that tend to be active when people are not engaged in an externally oriented task, and which appears to be crucial for thinking about experience that is not present to the senses, including episodic memory of the past, imagination of the future, and the imagined experience of others (Allen et al., 2017; Andrews-Hanna, Smallwood, & Spreng, 2014).

Most of the neural research relevant to Agreeableness has come from the study of empathy as a trait, which is a major component of Compassion. Individual differences in empathy have been linked to regions of the default network in both functional and structural MRI studies (Takeuchi et al., 2014a, 2014b). In addition to theory of mind, empathy is also related to the tendency to experience emotions vicariously, and this ability appears to be linked especially to the insula and ACC and to patterns of neural activity that are similar when one experiences distress oneself or witnesses someone else in distress. Several studies have found trait empathy to be positively related to the cortical volume or thickness of the insula (Mutschler, Reinbold, Wankerl, Seifritz, & Ball, 2013; Patil et al., 2017; Sassa et al., 2012; Valk et al., 2017), and another found Compassion (but not Politeness) to be positively correlated with cortical volume of the insula and ACC (Hou et al., 2017). Finally, Antagonism is positively associated with levels of the hormone testosterone, which appears to be most specifically linked to low Politeness, based on research relating testosterone to individual differences in interpersonal behavior and aggression (DeYoung et al., 2013; Montoya et al., 2012; Turan et al., 2014).

Psychoticism, Thought Disorder, and Openness to Experience

The final major dimension of psychopathology, Psychoticism or Thought Disorder, has been the most difficult to integrate with the Big Five model of personality variation (Kotov et al., 2017). Psychoticism reflects deficits in reality testing—that is, the tendency to make errors in interpreting the facts of the world one inhabits. Hallucinations and complex delusions are the extreme forms of these errors, but they simply represent extremes of a common

phenomenon in human mental life, which has been called *apophenia*, the detection of meaningful patterns where none in fact exist (DeYoung et al., 2012). Apophenia is equivalent to making Type I errors—false positives—in perception and belief. Everyday examples include seeing faces in inanimate objects, mistakenly hearing one's name, and holding superstitious beliefs, such as the gambler's fallacy or astrology. Studies in both healthy and clinical samples have demonstrated that the general tendency toward apophenia, including unusual perceptual experiences and magical ideation, is positively associated with the Openness aspect of the Openness/Intellect dimension, but at least weakly negatively related to the Intellect aspect (Chmielewski, Bagby, Markon, Ring, & Ryder, 2014; DeYoung et al., 2012, 2016). This differential association at the aspect level explains the lack of strong congruence at the Big Five level. Even given this confound, however, large molecular genetic studies have shown genetic correlations and overlapping genetic variants for Openness/Intellect and risk for schizophrenia (Lo et al., 2017; Smeland et al., 2017).

CB5T posits that Openness/Intellect reflects the capacity for cognitive exploration that allows people to generate interpretations of the past, present, and possible future in terms of their factual properties. Whereas Intellect reflects the ability and tendency to interpret the causal and logical structure of the world, through reasoning, Openness reflects the ability and tendency to interpret the correlational structure of the world, largely through perception and imagination (which is simulation of perception). Correlational structure refers broadly to anything that tends to co-occur, from basic patterns in visual perception to patterns of covariation among complex events. CB5T posits that Psychoticism stems primarily from high Openness but is exacerbated by the absence of sufficient intellectual capacity to weed out Type I errors in the multifarious patterns that are detected by highly open people (DeYoung, 2015). Psychoticism or apophenia can be considered *openness to highly implausible patterns* (DeYoung et al., 2012). This form of Openness is especially likely in conjunction with low levels of cognitive abilities, like intelligence and working memory capacity, that are associated with Intellect and that are reduced in the "disorganized" or "formal" symptoms of Thought Disorder (DeYoung et al., 2009, 2014; Goghari, Sponheim, & MacDonald, 2010; Kotov et al., 2016). Indeed, CB5T considers intelligence (which is well measured as IQ) to be a facet of Intellect (DeYoung, 2015; DeYoung et al., 2012), and we suggest that cognitive disorganization, inasmuch as it is distinct from Psychoticism, may be akin to intellectual disability and considered a disorder of Intellect, as depicted in Figure 2.

Human beings have evolved an astonishing capacity to create complex mental models of the world they inhabit, and many brain systems are likely to be involved in the vulnerability of those models to distortion. The neural correlates we review here are, thus, probably only a few among many. Both psychosis and Openness have been linked to dopamine (DeYoung, 2013; Howes, McCutcheon, Owen, & Murray, 2017; Maia & Frank, 2017). CB5T hypothesizes that Openness and Psychoticism are related to increased activity in a branch of the dopaminergic system distinct from that underlying Extraversion (Allen & DeYoung, 2017). Extraversion is linked to dopaminergic neurons that encode value, increasing their firing rate when outcomes are better than anticipated and decreasing it when worse than anticipated. Another set of dopaminergic neurons appears to encode salience, or the value of information regardless of valence; these neurons increase firing rate for both better- and worse-than-anticipated outcomes, project to different parts of the brain than value-coding neurons, and appear to facilitate cognitive exploration (Bromberg-Martin, Matsumoto, & Hikosaka, 2010; DeYoung, 2013). Functional connectivity analysis of fMRI data has provided evidence that Openness predicts synchronized activity between dopaminergic neurons in the midbrain and the DLPFC during pleasant sensory experiences and at rest (Passamonti et al., 2015). A process of "aberrant salience" may help to explain why people high in Openness and at risk for psychosis are prone to detecting patterns erroneously, if dopamine signals that valuable information is present even when that is highly unlikely (DeYoung et al., 2012; Kapur, 2003).

Another major neural correlate that has been well established for psychosis is altered white matter connectivity, especially reductions in white matter coherence in tracts connecting particular thalamocortical and frontotemporal regions (Cannon, 2015; Pettersson-Yeo, Allen, Benetti, McGuire, & Mechelli, 2011). Altered connectivity may contribute to atypical sensory and cognitive integration, leading to unusual modes of pattern detection. Openness has been found to predict reduced white matter coherence in the same frontal lobe tracts that are implicated in psychosis (Grazioplene, Chavez, Rustichini, & DeYoung, 2016; Jung, Grazioplene, Caprihan, Chavez, & Haier, 2010). Other neural processes involved in Psychoticism are also likely to show connections with Openness/Intellect. For example, psychosis has been linked to deficits in episodic memory, a cognitive function that depends on the default network (Cannon, 2015), and Openness/Intellect has been linked to functional parameters of that network (Beaty et al., 2016). In relation to cognitive disorganization, much is known about the neural correlates of intelligence and working memory, and questionnaire measures of Intellect share some of those correlates (DeYoung et al., 2009). The major network involved in these abilities is the fronto-parietal control network, which has also been implicated in cognitive disorganization (Goghari et al., 2010; Kotov et al., 2016).

Psychosis is considered one of the most severe forms of psychopathology, but we would argue that even hallucinations and delusions are not, by themselves, sufficient to indicate psychopathology. Imagine a woman who regularly hears voices and believes that she is communicating with the spirits of the dead—these are hallucinations and delusions. Now, imagine that this woman is a successful psychic with a flourishing business and a social network that either accepts or humors her beliefs, and she leads a relatively normal and happy life (Powers, Kelley, & Corlett, 2017). In the absence of cybernetic dysfunction, even psychotic symptoms are not

pathological. They are much more likely to be accompanied by psychopathology, we suspect, if the individual is unable to form normal social connections, a feature of the so-called negative symptoms of schizophrenia that are subsumed within Detachment in our theory.

The Metatraits and the General Factor of Psychopathology

One difference between measures of psychopathological traits or symptoms and normal personality measures is that the former show a greater tendency toward intercorrelation, such that there appears to be a general factor of psychopathology or *p*-factor (Caspi et al., 2014; Lahey, Krueger, Rathouz, Waldman, & Zald, 2017). In normal personality measures, the general factor is weaker and appears to be an artifact of the biases of individual raters, as it disappears when using scores from multiple raters (Chang, Connelly, & Geeza, 2012; DeYoung, 2006; McCrae et al., 2008; Revelle & Wilt, 2013). Hence, Figure 1 depicts the metatraits, Stability and Plasticity, as uncorrelated. An important question, then, is, What is the relation of the *p*-factor to the metatraits? Studies of both common mental disorders and personality disorders have shown that the *p*-factor is strongly related to Neuroticism but is additionally associated with low Conscientiousness and Agreeableness (Caspi et al., 2014; Castellanos-Ryan et al., 2016; Tackett et al., 2013; Wright, Hopwood, Skodol, & Morey, 2016). Thus, it strongly resembles Stability (DeYoung, 2006). The *p*-factor is also weakly negatively associated with IQ, as is Neuroticism, which may reflect that intelligence is somewhat protective against most forms of psychopathology or may simply be attributable to heightened test anxiety among people high in Neuroticism (Caspi et al., 2014; Castellanos-Ryan et al., 2016; DeYoung, 2011; Gale, Batty, Tynelius, Deary, & Rasmussen, 2010; Moutafi, Furnham, & Tsaousis, 2006).

The similarity between Stability and the general factor of psychopathology is consistent with our definition of psychopathology in terms of the failure of characteristic adaptations. CB5T describes the metatraits as reflecting variations in people's ability to meet the two most fundamental needs of any cybernetic system that can adapt to complex and changing environments: Stability is related to the need to maintain the stability of ongoing goal-directed functioning, whereas Plasticity is related to the need to engage in exploration to develop new adaptations. Stability reflects the tendency of existing characteristic adaptations to remain intact and to resist disruption by emotions, impulses, and doubts, whereas Plasticity reflects the tendency to generate new adaptations (DeYoung, 2015).

According to our theory, all psychopathology involves cybernetic dysfunction, failure to make progress toward important goals due to failure of characteristic adaptations. Breakdowns in the mechanisms associated with any particular dimension of psychopathology will tend to lead to dysfunction in the domain of behavior described by that dimension. Because the cybernetic system consists of interacting mechanisms, however, dysfunction in one mechanism is likely to lead to dysfunction in others—that is, in the system as a whole—and this general tendency toward dysfunction, captured by the *p*-factor, may be well described as an absence of Stability. Further, Stability may reflect not just an emergent property of the system's many interacting parts but also the operation of broadly acting mechanisms that evolved to regulate stability. Serotonin is a strong candidate as a factor directly influencing Stability (and hence also the *p*-factor) both because of what is known regarding serotonin's role in constraining impulses and facilitating goal-pursuit (Carver, Johnson, & Joormann, 2008; Gray & McNaughton, 2000; Spoont, 1992) and because variation in serotonergic function has been linked not only to Neuroticism but also to Agreeableness and Conscientiousness (Allen & DeYoung, 2017). Our characterization of the *p*-factor as low Stability is consistent with descriptions of both constructs in terms of inability to restrain urgent emotional impulses (Carver, Johnson, & Timpano, 2017; DeYoung & Rueter, 2016; Hirsh, DeYoung, & Peterson, 2009).

Whereas CB5T associates Stability with serotonin, it associates Plasticity with dopamine, which drives exploration both in behavior (Extraversion) and in cognition (Openness/Intellect; Allen & DeYoung, 2017; DeYoung, 2013). Lack of association of Plasticity with the *p*-factor may be due to the fact that Plasticity should facilitate mental health only inasmuch as it successfully increases Stability. Exploration is not always beneficial and can destabilize the system instead of generating effective and well-integrated characteristic adaptations. Indeed, Plasticity has been shown to be positively associated with externalizing problems in adolescent boys (DeYoung, Peterson, Séguin, Pihl, & Tremblay, 2008), and both Extraversion and Openness/Intellect are elevated in bipolar disorder (Tackett et al., 2008). Nonetheless, some amount of exploration is necessary for stable cybernetic function in any sufficiently unpredictable environment. Thus, having a greater capacity to generate new characteristic adaptations may be protective against many forms of psychopathology, given that our definition requires not only that characteristic adaptations fail but also that they are not soon replaced with effective new ones. Relations among the metatraits, the *p*-factor, and the tendency toward resilience is an intriguing topic for future research.

Conclusion: Implications for Research, Diagnosis, and Intervention

Having articulated our cybernetic theory of psychopathology, we conclude with a relatively brief consideration of its implications. First, research on the etiology of psychopathology should move from studying diagnostic categories to studying continuous dimensions of psychopathology. The best research designs will involve community or treatment-seeking samples, rather than being selected on the basis of diagnosis as in case-control designs. This means ignoring current diagnostic categories and focusing on "transdiagnostic" constructs in a manner similar to the National Institute of Mental Health's Research Domain

Criteria (RDoC) initiative. The five major dimensions of psychopathology are reasonably similar to the domains of RDoC (although we believe that their Cognitive Systems domain, targeting executive deficits, could usefully be fractionated into two separate domains related to Disinhibition and Thought Disorder) and can fruitfully be integrated into RDoC-oriented research. Such research should not limit itself to studying the five major dimensions but should also consider their subdivisions into aspects and facets.

Our theory posits that the isomorphism of the major dimensions of psychopathology with normal personality traits is because the same mechanisms that underlie normal personality are precisely those that produce dysfunction in most cases of psychopathology, due to unusually high or low levels of function, either chronically or in a temporary fluctuation away from one's typical level. This isomorphism implies that research in psychopathology can be facilitated by studying normal personality variation in parallel. Our discussion of neural variables linked to dimensions of personality and psychopathology was not intended to be comprehensive, but, along with our identification of cybernetic functions associated with each dimension, it provides a wealth of potential hypotheses for future research. According to our theory, what are typically considered as symptoms or consequences of psychopathology are probably more often contributing causes of psychopathology. People are unable to pursue their goals and develop successful new adaptations because extreme levels of certain traits render the cybernetic system unstable and inflexible. Nonetheless, given that our theory specifies that extreme trait levels are not, by themselves, sufficient to identify psychopathology, a particularly important area for future research will be investigating what factors increase the likelihood that an extreme trait leads to dysfunction. It may also be fruitful to conduct research on the dynamics through which common types of characteristic *mal*adaptation lead to dysfunction even in the absence of extreme trait levels.

Not only should the current diagnostic categories be avoided in research, we believe they should be replaced in official diagnostic systems as well. Most existing diagnoses are so heterogeneous that they are ambiguous in identifying likely sources of psychopathology and, hence, ambiguous in directing treatment (Clark, Watson, & Reynolds, 1995; Harkness et al., 2014; Hasler, Drevets, Manji, & Charney, 2004; Zimmerman, Ellison, Young, Chelminski, & Dalrymple, 2015). Current diagnoses neither accurately represent which symptoms tend to co-occur within individuals nor reflect conditions with functionally coherent etiologies. We argue that the actual causal processes, which should serve as loci of control for prevention and treatment, are often associated with traits that can be located within the Big Five personality hierarchy. A growing movement, exemplified by the Hierarchical Taxonomy of Psychopathology consortium, is attempting to replace current nosologies with empirically based models like the one we propose here (Kotov et al., 2017). To get from our theory of psychopathology to a working diagnostic system will require a process of sociopolitical negotiation to determine what level of severity is worthy of diagnosis in which contexts, but this is fundamentally no different than the process by which it is determined, for example, what level of insulin response is worthy of the diagnosis "type 2 diabetes."

In the improved clinical approaches that we envision, diagnosis should proceed using the criterion of cybernetic dysfunction—that is, determination that important goals are unable to be achieved and effective new characteristic adaptations have persistently failed to be developed. Once identified, dysfunction should additionally be characterized according to the dimensions of psychopathology that the individual exhibits, at multiple levels of the trait hierarchy. Reviewing mental health in terms of the functions associated with each of the major trait dimensions is analogous to a "review of systems" approach in physical health (Harkness et al., 2014). When people seek treatment for reasons of mental health, it is sensible to evaluate how they are functioning in all major areas of psychological function, regardless of their presenting complaints.

Our perspective leads to multiple conclusions regarding intervention. First, it lends itself to a dimensional matching of severity to treatment. There need not be only a single diagnostic threshold, even for dysfunction within a single trait dimension, and different degrees of psychopathology may be amenable to different treatments. Second, clinicians should pay attention to whether the cause of psychopathology in a given case is more likely to be due to extreme trait levels or to some particularly problematic characteristic adaptations, as this distinction may have implications for the optimal locus of control. Figure 3 illustrates the causal dynamics that CB5T identifies in both the sources and the consequences of personality. If the cause of psychopathology appears to be that extremes of some trait(s) are chronically leading to the failure of characteristic adaptations, then one may wish to intervene with pharmaceuticals, thereby attempting to adjust directly the cybernetic mechanisms underlying the trait(s) in question. However, it may be equally or even more effective to target characteristic adaptations directly, through behavioral or talk therapies.

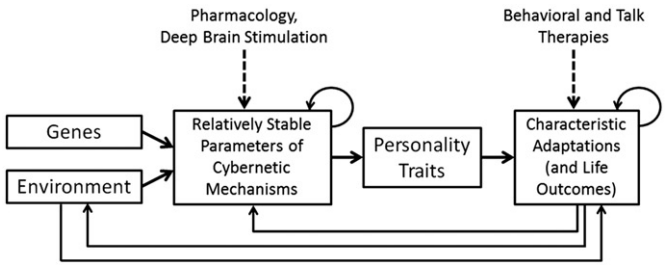

Figure 3. Causal processes in the functioning of personality and treatment of psychopathology. *Note.* Solid arrows represent spontaneous causal processes. Dashed arrows represent therapeutic interventions. Both genetics and the environment directly influence the cybernetic mechanisms underlying personality traits. All genetic influences on characteristic adaptations are funneled through traits, but the environment can influence characteristic adaptations independently of the influence of traits. Circular arrows indicate that cybernetic parameters can influence each other, as can characteristic adaptations and other life outcomes, such as physical health. Neurobiological interventions are aimed directly at changing the parameters of cybernetic mechanisms that cause variation in personality traits. Behavioral or talk therapy interventions are aimed directly at changing characteristic adaptations. Each kind of intervention can influence the other component of personality indirectly.

Such therapies can help people to interpret their experience in different ways, to develop new strategies for pursuing their goals, and to change their goals, either by adding and subtracting goals or by rethinking the importance of particular goals.

One of the most important features of Figure 3 is that the causal paths between traits and characteristic adaptations flow in both directions, as characteristic adaptations can feed back to change the cybernetic parameters underlying traits. Someone who learns various strategies to act in a more conscientious manner, for example, such as using a structured to-do list and dedicated workspaces, may become better able to prioritize goals effectively and resist distractions in general, thus increasing in trait Conscientiousness. Voluntary changes in characteristic adaptations can lead to trait change (Hennecke, Bleidorn, Denissen, & Wood, 2014; Hudson & Fraley, 2015; Hudson, Roberts, & Lodi-Smith, 2012).

For three reasons, we believe that interventions targeting characteristic adaptations directly are often likely to be more effective than pharmacological interventions and to have fewer unwanted side effects. First, extreme trait levels are neither necessary nor sufficient for psychopathology. Second, failure of characteristic adaptations to allow successful goal pursuit is the core of our definition of psychopathology, such that developing effective new characteristic adaptations is ultimately the only route to mental health, even if the intervention initially targets traits. Third, the neural systems underlying any given trait are highly complex, and our understanding of those systems is currently limited, making it difficult to target the optimal neurobiological parameters to adjust problematic traits.

A recent meta-analysis of the effects of clinical interventions on personality traits supports our position. All methods of intervention that were examined produced significant trait change, but trait changes due to pharmacological interventions were smaller in magnitude than those due to cognitive behavioral therapy and supportive or humanistic therapy (though closer in magnitude to those due to psychodynamic therapy; Roberts et al., 2017). Consistent with our proposition that Neuroticism is a common consequence (as well as cause) of psychopathology, Neuroticism is the trait that changed most following all kinds of intervention. As people develop new characteristic adaptations that let them proceed toward their goals more effectively, they experience less severe mismatches between their desired state and their current situation, and hence they have fewer defensive reactions to mismatch and experience less negative emotion.

In relation to the prevention of psychopathology, our theory suggests the utility of personality-targeted interventions (O'Leary-Barrett, Castellanos-Ryan, Pihl, & Conrod, 2016). Such interventions attempt to teach new characteristic adaptations that are matched to each person's profile of risky personality traits. These characteristic adaptations can help to offset the particular vulnerabilities associated with the relevant traits. Personality-targeted interventions have proven to be impressively successful in reducing substance use and mental health problems in adolescents (Conrod, Castellanos-Ryan, & Strang, 2010; Newton et al., 2016).

Our cybernetic theory of psychopathology offers a new perspective on research, diagnosis, and intervention. We believe it provides a more coherent and useful definition of psychopathology than any existing definition. It also provides a mechanistic account of the causes and manifestations of psychopathology, though this account is naturally incomplete, as much remains to be learned at both psychological and neurobiological levels. Nonetheless, the theory provides a framework in which the most important classes of function are described and in which new details can be integrated in a way that leads to a cohesive body of knowledge and theory rather than a fragmented list of associations between poorly defined diagnostic categories and psychological or neural parameters that lack clear functional connections with one another or with the psychopathology they are intended to explain. The cybernetic perspective allows a truly systematic and synthetic understanding of psychopathology.

References

Abram, S. V., Wisner, K. M., Grazioplene, R. G., Krueger, R. F., MacDonald, A. W., & DeYoung, C. G. (2015). Functional coherence of insula networks is associated with externalizing behavior. *Journal of Abnormal Psychology*, *124*, 1079–1091. doi:10.1037/abn0000078

Allen, T. A., & DeYoung, C. G. (2017). Personality neuroscience and the Five Factor Model. In T. A. Widiger (Ed.). *Oxford handbook of the Five Factor Model* (pp. 319–349). New York, NY: Oxford University Press.

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,* 599–613. doi: 10.1002/per.2133

Andrews, P. W., Bharwani, A., Lee, K. R., Fox, M., & Thomson, J. A. (2015). Is serotonin an upper or a downer? The evolution of the serotonergic system and its role in depression and the antidepressant response. *Neuroscience and biobehavioral reviews*, *51*, 164–188. doi: 10.1016/j.neubiorev.2015.01.018

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*, 29–52. doi:10.1111/nyas.12360

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*, 338–375. doi:10.1037/0033-2909.120.3.338

Beaty, R. E., Kaufman, S. B., Benedek, M., Jung, R. E., Kenett, Y. N., Jauk, E., … Silvia, P. J. (2016). Personality and complex brain networks: The role of openness to experience in default network efficiency. *Human Brain Mapping*, *37*, 773–779. doi:10.1002/hbm.23065

Bechhoefer, J. (2005). Feedback for physicists: A tutorial essay on control. *Reviews of Modern Physics*, *77*, 783–836. doi:10.1103/RevModPhys.77.783

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: 'Liking', 'wanting', and learning. *Current Opinion in Pharmacology*, *9*, 65–73. doi:10.1016/j.coph.2008.12.014

Bleidorn, W., Hopwood, C. J., & Lucas, R. E. (2018). Life events and personality trait change. *Journal of Personality*, *86*, 83–96. doi: 10.1111/jopy.12286

Bolton, D. (2008). *What is mental disorder? An essay in philosophy, science, and values*. New York, NY: Oxford University Press.

Boorse, C. (1977). Health as a theoretical concept. *Philosophy of science*, *44*, 542–573. doi:10.1086/288768

Boorse, C. (2011). Concepts of health and disease. In F. Gifford (Ed.), *Handbook of the Philosophy of Science, vol. 16: Philosophy of Medicine* (pp. 13–64). New York: Elsevier.

Bress, J. N., Smith, E., Foti, D., Klein, D. N., & Hajcak, G. (2012). Neural response to reward and depressive symptoms in late childhood to early adolescence. *Biological Psychology*, *89*, 156–162. doi:10.1016/j.biopsycho.2011.10.004

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*, 815–834. doi:10.1016/j.neuron.2010.11.022

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, *15*, 118–121. doi:10.1111/j.0963-7214.2006.00419.x

Cannon, T. D. (2015). How schizophrenia develops: cognitive and brain mechanisms underlying onset of psychosis. *Trends in cognitive sciences*, *19*, 744–756. doi:10.1016/j.tics.2015.09.009

Carragher, N., Krueger, R. F., Eaton, N. R., Markon, K. E., Keyes, K. M., Blanco, C., … Hasin, D. S. (2014). ADHD and the externalizing spectrum: Direct comparison of categorical, continuous, and hybrid models of liability in a nationally representative sample. *Social Psychiatry and Psychiatric Epidemiology*, *49*, 1307–1317. doi:10.1007/s00127-013-0770-3

Carver, C. S., Johnson, S. L., & Joormann, J. (2008). Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: what depression has in common with impulsive aggression. *Psychological Bulletin*, *134*, 912–943. doi:10.1037/a0013740

Carver, C. S., Johnson, S. L., & Timpano, K. R. (2017). Toward a Functional View of the p Factor in Psychopathology. *Clinical Psychological Science*. Advance online publication. doi:10.1177/2167702617710037

Carver, C., & Scheier, M. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press.

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, HLee., Israel, S., … Moffitt, T. E. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*, 119–137. doi:10.1177/2167702613497473

Caspi, A., & Shiner, R. L. (2006). Personality development. In N. Eisenberg (Ed.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development* (6th ed., pp. 300–365). New York, NY: Wiley.

Castellanos-Ryan, N., Brière, F. N., O'Leary-Barrett, M., Banaschewski, T., Bokde, A., Bromberg, U., … Garavan, H. (2016). The structure of psychopathology in adolescence and its common personality and cognitive correlates. *Journal of Abnormal Psychology*, *125*, 1039–1052. doi:10.1037/abn0000193

Chang, L., Connelly, B. S., & Geeza, A. A. (2012). Separating method factors and higher order traits of the Big Five: A meta-analytic multitrait-multimethod approach. *Journal of Personality and Social Psychology*, *102*, 408–426. doi:10.1037/a0025559

Cheetham, A., Allen, N. B., Whittle, S., Simmons, J., Yücel, M., & Lubman, D. I. (2017). Orbitofrontal cortex volume and effortful control as prospective risk factors for substance use disorder in adolescence. *European Addiction Research*, *23*, 37–44. doi:10.1159/000452159

Chmielewski, M., Bagby, R. M., Markon, K., Ring, A. J., & Ryder, A. G. (2014). Openness to experience, intellect, schizotypal personality disorder, and psychoticism: Resolving the controversy. *Journal of Personality Disorders*, *28*, 483–499. doi:10.1521/pedi_2014_28_128

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181–204. doi:10.1017/S0140525X12000477

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

Conrod, P. J., Castellanos-Ryan, N., & Strang, J. (2010). Brief, personality-targeted coping skills interventions and survival as a non-drug user over a 2-year period during adolescence. *Archives of General Psychiatry*, *67*, 85–93. doi:10.1001/archgenpsychiatry.2009.173

Del Giudice, M. (2015). Self-regulation in an evolutionary perspective. In G. H. E. Gendolla, M. Tops, & S. Koole (Eds.), *Handbook of Biobehavioral Approaches to Self-Regulation* (pp. 25–41). New York, NY: Springer. doi:10.1007/978-1-4939-1236-0_3

Del Giudice, M. (2016). The life history model of psychopathology explains the structure of psychiatric disorders and the emergence of the p factor: a simulation study. *Clinical Psychological Science*, *4*, 299–311. doi:10.1177/2167702615583628

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*, 491–569. doi:10.1017/S0140525X99002046

Depue, R. A., & Morrone-Strupinsky, J. V. (2005). A neurobehavioral model of affiliative bonding: Implications for conceptualizing a human trait of affiliation. *Behavioral and Brain Sciences*, *28*, 313–350. doi:10.1017/S0140525X05000063

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, *91*, 1138–1151. doi:10.1037/0022-3514.91.6.1138

DeYoung, C. G. (2011). Intelligence and personality. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 711–737). New York: Cambridge University Press.

DeYoung, C. G. (2013). The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. *Frontiers in Human Neuroscience*, *7* article 762. doi:10.3389/fnhum.2013.00762

DeYoung, C. G. (2015). Cybernetic Big Five Theory. *Journal of Research in Personality*, *56*, 33–58. doi:10.1016/j.jrp.2014.07.004

DeYoung, C. G., Carey, B. E., Krueger, R. F., & Ross, S. R. (2016). Ten aspects of the Big Five in the personality inventory for DSM–5. *Personality Disorders: Theory, Research, and Treatment*, *7*, 113–123. doi:10.1037/per0000170

DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality*, *46*, 63–78. doi:10.1016/j.jrp.2011.12.003

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of Openness/Intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*, 825–858. doi:10.1111/j.1467-6494.2005.00330.x

DeYoung, C. G., Peterson, J. B., Séguin, J. R., Pihl, R. O., & Tremblay, R. E. (2008). Externalizing behavior and the higher-order factors of the Big Five. *Journal of Abnormal Psychology*, *117*, 947–953. doi:10.1037/a0013742

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*, 880–896. doi:10.1037/0022-3514.93.5.880

DeYoung, C. G., Quilty, L. C., Peterson, J. B., & Gray, J. R. (2014). Openness to experience, intellect, and cognitive ability. *Journal of Personality Assessment*, *96*, 46–52. doi:10.1080/00223891.2013.806327

DeYoung, C. G., & Rueter, A. R. (2016). Impulsivity as a personality trait. In K. D. Vohs & R. F. Baumeister (Ed.). *Handbook of self-regulation: research, theory, and applications* (3rd ed., pp. 345–363). New York, NY: Guilford Press.

DeYoung, C. G., Shamosh, N. A., Green, A. E., Braver, T. S., & Gray, J. R. (2009). Intellect as distinct from Openness: Differences revealed by fMRI of working memory. *Journal of Personality and Social Psychology*, *97*, 883–892. doi:10.1037/a0016615

DeYoung, C. G., & Weisberg, Y. J. (2018). Cybernetic approaches to personality and social behavior. In M. Snyder & K. Deaux (Ed.), *Oxford Handbook of Personality and Social Psychology* (2nd ed., pp. 387–414). New York, NY: Oxford University Press.

DeYoung, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the aspects of the Big Five, the interpersonal circumplex, and trait affiliation. *Journal of Personality*, *81*, 465–475. doi:10.1111/jopy.12020

Di Simplicio, M., Norbury, R., Reinecke, A., & Harmer, C. J. (2014). Paradoxical effects of short-term antidepressant treatment in fMRI

emotional processing models in volunteers with high neuroticism. *Psychological Medicine*, *44*, 241–252. doi:10.1017/S0033291713000731

Du, L., Bakish, D., Ravindran, A. V., & Hrdina, P. D. (2002). Does fluoxetine influence major depression by modifying five-factor personality traits? *Journal of Affective Disorders*, *71*(1–3), 235–241. doi:10.1016/S0165-0327(01)00370-6

Elliot, A. J., & Fryer, J. W. (2008). The goal construct in psychology. In J. Y. Shah & W. L. Gardner (Eds.), *Handbook of motivation science* (pp. 235–250). New York, NY: Guilford.

Finnigan, K. M., & Vazire, S. (2018). The incremental validity of average state self-reports over global self-reports of personality. *Journal of Personality and Social Psychology*, *115*, 321–337. doi:10.1037/pspp0000136

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, *97*, 1097–1114. doi:10.1037/a0016786

Forbes, M. K., Kotov, R., Ruggero, C. J., Watson, D., Zimmerman, M., & Krueger, R. F. (2017). Delineating the joint hierarchical structure of clinical and personality disorders in an outpatient psychiatric sample. *Comprehensive Psychiatry*, *79*, *19-30*. doi:10.1016/j.comppsych.2017.04.006

Friston, K. J. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. doi:10.1038/nrn2787

Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science*, *2*, 31–39. doi:10.1111/j.1467-9280.1991.tb00093.x

Gale, C. R., Batty, G. D., Tynelius, P., Deary, I. J., & Rasmussen, F. (2010). Intelligence in early adulthood and subsequent hospitalization for mental disorders. *Epidemiology*, *21*, 70–77. doi:10.1097/EDE.0b013e3181c17da8

Goghari, V. M., Sponheim, S. R., & MacDonald, A. W. (2010). The functional neuroanatomy of symptom dimensions in schizophrenia: A qualitative and quantitative review of a persistent question. *Neuroscience & Biobehavioral Reviews*, *34*, 468–486. doi:10.1016/j.neubiorev.2009.09.004

Gore, W. L., & Widiger, T. A. (2013). The DSM-5 dimensional trait model and five-factor models of general personality. *Journal of Abnormal Psychology*, *122*, 816–821. doi:10.1037/a0032822

Gray, J. A. (1973). Causal theories of personality and to test them. In Royce, J. R. (Ed.), *Multivariate analysis and psychological theory* (pp. 409–463). New York, NY: Academic Press.

Gray, J. A. (2004). *Consciousness: Creeping up on the hard problem*. New York, NY: Oxford University Press. doi:10.1002/acp.1121

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system* (2nd ed.). New York, NY: Oxford University Press.

Grazioplene, R. G., Chavez, R. S., Rustichini, A., & DeYoung, C. G. (2016). White matter correlates of psychosis-linked traits support continuity between personality and psychopathology. *Journal of Abnormal Psychology*, *125*, 1135–1145. doi:10.1037/abn0000176

Griffith, J. W., Zinbarg, R. E., Craske, M. G., Mineka, S., Rose, R. D., Waters, A. M., & Sutton, J. M. (2010). Neuroticism as a common dimension in the internalizing disorders. *Psychological Medicine*, *40*, 1125–1136. doi:10.1017/S0033291709991449

Hall, J. R., & Benning, S. D. (2006). The "successful" psychopath: Adaptive and subclinical manifestations of psychopathy in the general population. In C. J. Patrick (Ed.), *Handbook of Psychopathy* (pp. 459–478). New York, NY: Guilford.

Harkness, A. R., Reynolds, S. M., & Lilienfeld, S. O. (2014). A review of systems for psychology and psychiatry: Adaptive systems, Personality Psychopathology Five (PSY–5), and the DSM–5. *Journal of Personality Assessment*, *96*, 121–139. doi:10.1080/00223891.2013.823438

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

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

Hennecke, M., Bleidorn, W., Denissen, J. J., & Wood, D. (2014). A three-part framework for self-regulated personality development across adulthood. *European Journal of Personality*, *28*, 289–299. doi:10.1002/per.1945

Henry, J., Pingault, J. B., Boivin, M., Rijsdijk, F., & Viding, E. (2016). Genetic and environmental aetiology of the dimensions of Callous-Unemotional traits. *Psychological Medicine*, *46*, 405–414. doi:10.1017/S0033291715001919

Hirsh, J. B., DeYoung, C. G., & Peterson, J. B. (2009). Metatraits of the Big Five differentially predict engagement and restraint of behavior. *Journal of Personality*, *77*, 1085–1102. doi:10.1111/j.1467-6494.2009.00575.x

Hirsh, J. B., Mar, R. A., & Peterson, J. B. (2012). Psychological entropy: A framework for understanding uncertainty-related anxiety. *Psychological Review*, *119*, 304-320. doi:10.1037/a0026767

Holmes, A. J., Lee, P. H., Hollinshead, M. O., Bakst, L., Roffman, J. L., Smoller, J. W., & Buckner, R. L. (2012). Individual differences in amygdala-medial prefrontal anatomy link negative affect, impaired social functioning, and polygenic depression risk. *The Journal of Neuroscience*, *32*, 18087–18100. doi:10.1523/JNEUROSCI.2531-12.2012

Hou, X., Allen, T. A., Wei, D., Huang, H., Wang, K., DeYoung, C. G., & Qiu, J. (2017). Trait compassion is associated with the neural substrate of empathy. *Cognitive, Affective, & Behavioral Neuroscience*, *17*, 1018–1027. doi:10.3758/s13415-017-0529-5

Howes, O. D., McCutcheon, R., Owen, M. J., & Murray, R. M. (2017). The role of genes, stress, and dopamine in the development of schizophrenia. *Biological Psychiatry*, *81*, 9–20. doi:10.1016/j.biopsych.2016.07.014

Hudson, N. W., & Fraley, R. C. (2015). Volitional personality trait change: Can people choose to change their personality traits? *Journal of Personality and Social Psychology*, *109*, 490–507. doi:10.1037/pspp0000021

Hudson, N. W., Roberts, B. W., & Lodi-Smith, J. (2012). Personality trait development and social investment in work. *Journal of Research in Personality*, *46*, 334–344. doi:10.1016/j.jrp.2012.03.002

Ilieva, I. (2015). Enhancement of healthy personality through psychiatric medication: The influence of SSRIs on neuroticism and extraversion. *Neuroethics*, *8*, 127–137. doi:10.1007/s12152-014-9226-z

Jackson, J., Balota, D. A., & Head, D. (2011). Exploring the relationship between personality and regional brain volume in healthy aging. *Neurobiology of Aging*, *32*, 2162–2171. doi:10.1016/j.neurobiolaging.2009.12.009

Jang, K. L., Hu, S., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Personality and Individual Differences*, *33*, 83–101. doi:10.1016/S0191-8869(01)00137-4

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of personality. *Journal of Personality and Social Psychology*, *74*, 1556–1565. doi:10.1037/0022-3514.74.6.1556

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History: measurement, and conceptual issue. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.). *Handbook of personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.

Jung, R. E., Grazioplene, R., Caprihan, A., Chavez, R. S., & Haier, R. J. (2010). White matter integrity, creativity, and psychopathology: Disentangling constructs with diffusion tensor imaging. *PloS One*, *5*, e9818. doi:10.1371/journal.pone.0009818

Kamarck, T. W., Haskett, R. F., Muldoon, M., Flory, J. D., Anderson, B., Bies, R., … Manuck, S. B. (2009). Citalopram intervention for hostility: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *77*, 174–188. doi:10.1037/a0014394

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *The American Journal of Psychiatry*, *160*, 13–23. doi:10.1176/appi.ajp.160.1.13

Kendler, K. S., & Schaffner, K. F. (2011). The dopamine hypothesis of schizophrenia: An historical and philosophical analysis. *Philosophy, Psychiatry, & Psychology*, *18*, 41–63. doi:10.1353/ppp.2011.0005

Kotov, R., Foti, D., Li, K., Bromet, E. J., Hajcak, G., & Ruggero, C. J. (2016). Validating dimensions of psychosis symptomatology: Neural correlates and 20-year outcomes. *Journal of Abnormal Psychology*, *125*, 1103–1119. doi:10.1037/abn0000188

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

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., … Eaton, N. R. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Aabnormal Psychology*, *126*, 454–478. doi:10.1037/abn0000258

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*, 1879–1890. doi:10.1017/S0033291711002674

Krueger, R. F., & Markon, K. E. (2014). The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology*, *10*, 477–501. doi:10.1146/annurev-clinpsy-032813-153732

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology*, *116*, 645–666. doi:10.1037/0021-843X.116.4.645

Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, *143*, 142–186. doi:10.1037/bul0000069

Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology*, *104*, 411–420. doi:10.1037/0021-843X.104.3.411

Lilienfeld, S. O., & Marino, L. (1999). Essentialism revisited: Evolutionary theory and the concept of mental disorder. *Journal of Abnormal Psychology*, *108*, 400–411. doi:10.1037/0021-843X.108.3.400

Lo, M. T., Hinds, D. A., Tung, J. Y., Franz, C., Fan, C. C., Wang, Y., … Sanyal, N. (2017). Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nature Genetics*, *49*, 152–156. doi:10.1038/ng.3736

Maia, T. V., & Frank, M. J. (2017). An integrative perspective on the role of dopamine in schizophrenia. *Biological Psychiatry*, *81*, 52–66. doi:10.1016/j.biopsych.2016.05.021

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., … Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*, 747–753. doi:10.1038/nature08494

Mansell, W. (2005). Control theory and psychopathology: An integrative approach. *Psychology and Psychotherapy: Theory, Research and Practice*, *78*, 141–178. doi:10.1348/147608304X21400

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*, 856–879. doi:10.1037/a0025727

Markon, K. E., & Krueger, R. F. (2005). Categorical and continuous models of liability to externalizing disorders: A direct comparison in NESARC. *Archives of General Psychiatry*, *62*, 1352–1359. doi:10.1001/archpsyc.62.12.1352

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139–157. doi:10.1037/0022-3514.88.1.139

Matsuo, K., Nicoletti, M., Nemoto, K., Hatch, J. P., Peluso, M. A., Nery, F. G., & Soares, J. C. (2009). A voxel-based morphometry study of frontal gray matter correlates of impulsivity. *Human Brain Mapping*, *30*, 1188–1195. doi:10.1002/hbm.20588

McCrae, R. R., Jang, K. L., Ando, J., Ono, Y., Yamagata, S., Riemann, R., … Spinath, F. M. (2008). Substance and artifact in the higher-order factors of the big five. *Journal of Personality and Social Psychology*, *95*, 442–455. doi:10.1037/0022-3514.95.2.442

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167

Miller, J. D., & Lynam, D. R. (2003). Psychopathy and the five-factor model of personality: A replication and extension. *Journal of Personality Assessment*, *81*, 168–178. doi:10.1207/S15327752JPA8102_08

Montoya, E. R., Terburg, D., Bos, P. A., & Van Honk, J. (2012). Testosterone, cortisol, and serotonin as key regulators of social aggression: A review and theoretical perspective. *Motivation and Emotion*, *36*, 65–73. doi:10.1007/s11031-011-9264-3

Moutafi, J., Furnham, A., & Tsaousis, I. (2006). Is the relationship between intelligence and trait neuroticism mediated by test anxiety? *Personality and Individual Differences*, *40*, 587–597. doi:10.1016/j.paid.2005.08.004

Munafò, M. R., & Flint, J. (2011). Dissecting the genetic architecture of human personality. *Trends in Cognitive Sciences*, *15*, 395–400. doi:10.1016/j.tics.2011.07.007

Mutschler, I., Reinbold, C., Wankerl, J., Seifritz, E., & Ball, T. (2013). Structural basis of empathy and the domain general region in the anterior insular cortex. *Frontiers in Human Neuroscience*, *7*, 177. doi:10.3389/fnhum.2013.00177

Nettle, D., & Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *European Journal of Personality*, *22*, 323–335. doi:10.1002/per.672

Newton, N. C., Conrod, P. J., Slade, T., Carragher, N., Champion, K. E., Barrett, E. L., … Teesson, M. (2016). The long-term effectiveness of a selective, personality-targeted prevention program in reducing alcohol use and related harms: a cluster randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *57*, 1056–1065. doi:10.1111/jcpp.12558

Nouchi, R., Takeuchi, H., Taki, Y., Sekiguchi, A., Kotozaki, Y., Nakagawa, S., … Kawashima, R. (2016). Neuroanatomical bases of effortful control: evidence from a large sample of young healthy adults using voxel-based morphometry. *Scientific Reports*, *6*, 31231. doi:10.1038/srep31231

O'Leary-Barrett, M., Castellanos-Ryan, N., Pihl, R. O., & Conrod, P. J. (2016). Mechanisms of personality-targeted intervention effects on adolescent alcohol misuse, internalizing and externalizing symptoms. *Journal of Consulting and Clinical Psychology*, *84*, 438–452. doi:10.1037/ccp0000082

Passamonti, L., Terracciano, A., Riccelli, R., Donzuso, G., Cerasa, A., Vaccaro, M. G., … Quattrone, A. (2015). Increased functional connectivity within mesocortical networks in open people. *NeuroImage*, *104*, 301–309. doi:10.1016/j.neuroimage.2014.09.017

Patil, I., Zanon, M., Novembre, G., Zangrando, N., Chittaro, L., & Silani, G. (2017). Neuroanatomical basis of concern-based altruism in virtual environment. *Neuropsychologia*, *116*, 34–43. doi:10.1016/j.neuropsychologia.2017.02.015

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*, 913–938. doi:10.1017/S0954579409000492

Peterson, J. B., & Flanders, J. (2002). Complexity management theory: Motivation for ideological rigidity and social conflict. *Cortex*, *38*, 429–458. doi:10.1016/S0010-9452(08)70680-4

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., & Mechelli, A. (2011). Dysconnectivity in schizophrenia: Where are we now? *Neuroscience and Biobehavioral Reviews*, *35*, 1110–1124. doi:10.1016/j.neubiorev.2010.11.004

Pickering, A. D. (2004). The neuropsychology of impulsive antisocial sensation seeking personality traits: From dopamine to hippocampal function? In R. M. Stelmack (Ed.), *On the psychobiology of personality: Essays in honor of Marvin Zuckerman* (pp. 453–477). New York, NY: Elsevier.

Pickering, A. D., & Gray, J. A. (1999). The neuroscience of personality. In L. Pervin & O. John (Eds.), *Handbook of personality*. (2nd ed., pp. 277–299). New York, NY: Guilford Press.

Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., … Fava, M. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *American Journal of Psychiatry*, *166*, 702–710. doi:10.1176/appi.ajp.2008.08081201

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*, 76–87. doi:10.1016/j.jpsychires.2008.03.001

Powers, W. T. (1973). *Behaviour: The control of perception*. Chicago, IL: Aldine.

Powers, A. R., Kelley, M. S., & Corlett, P. R. (2017). Varieties of voice-hearing: Psychics and the psychosis continuum. *Schizophrenia Bulletin*, *43*, 84–98. doi:10.1093/schbul/sbw133

Quilty, L. C., Meusel, L. A. C., & Bagby, R. M. (2008). Neuroticism as a mediator of treatment response to SSRIs in major depressive disorder. *Journal of Affective Disorders*, *111*, 67–73. doi:10.1016/j.jad.2008.02.006

Quilty, L. C., Pelletier, M., DeYoung, C. G., & Bagby, R. M. (2013). Hierarchical personality traits and the distinction between unipolar and bipolar disorders. *Journal of Affective Disorders*, *147*(1–3), 247–254. doi:10.1016/j.jad.2012.11.012

Revelle, W., & Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, *47*, 493–504. doi:10.1016/j.jrp.2013.04.012

Riccelli, R., Toschi, N., Riccelli, R., Toschi, N., Nigro, S., Terracciano, A., & Passamonti, L. (2017). Surface-based morphometry reveals the neuroanatomical basis of the five-factor model of personality. *Social Cognitive and Affective Neuroscience*, *12*, 671–684. doi:10.1093/scan/nsw175

Richters, J. E., & Hinshaw, S. P. (1999). The abduction of disorder in psychiatry. *Journal of Abnormal Psychology*, *108*, 438–445. doi:10.1037/0021-843X.108.3.438

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, *143*, 117–141. doi:10.1037/bul0000088

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 375–398). New York, NY: Guilford Press.

Robinson, M. D., Moeller, S. K., & Ode, S. (2010). Extraversion and reward-related processing: Probing incentive motivation in affective priming tasks. *Emotion*, *10*, 615. doi:10.1037/a0019173

Rueter, A. R., Abram, S. V., MacDonald, A. W., Rustichini, A., & DeYoung, C. G. (2018). The goal priority network as a neural substrate of Conscientiousness. *Human Brain Mapping*, *39*, 3574–3585. doi:10.1002/hbm.24195

Rudebeck, P. H., & Murray, E. A. (2011). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *Journal of Neuroscience*, *31*, 10569–10578. doi:10.1523/JNEUROSCI.0091-11.2011

Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review*, *28*, 1326–1342. doi:10.1016/j.cpr.2008.07.002

Sassa, Y., Taki, Y., Takeuchi, H., Hashizume, H., Asano, M., Asano, K., … Kawashima, R. (2012). The correlation between brain gray matter volume and empathizing and systemizing quotients in healthy children. *NeuroImage*, *60*, 2035–2041. doi:10.1016/j.neuroimage.2012.02.021

Schrödinger, E. (1944/1967). *What is life*? New York: Cambridge University Press.

Schwartenbeck, P., FitzGerald, T., Dolan, R. J., & Friston, K. (2013). Exploration, novelty, surprise, and free energy minimization. *Frontiers in Psychology*, *4*, 710. doi:10.3389%2Ffpsyg.2013.00710

Shackman, A. J., Tromp, D. P., Stockbridge, M. D., Kaplan, C. M., Tillman, R. M., & Fox, A. S. (2016). Dispositional negativity: An integrative psychological and neurobiological perspective. *Psychological Bulletin*, *142*, 1275–1314. doi:10.1037/bul0000073

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 623–423, 623–656. doi:10.1002/j.1538-7305.1948.tb01338.x

Smeland, O. B., Wang, Y., Lo, M.-T., Li, W., Frei, O., Witoelar, A., … Andreassen, O. A. (2017). Identification of genetic loci shared between schizophrenia and the Big Five personality traits. *Scientific Reports*, *7*, 2222. doi:10.1038/s41598-017-02346-3

Smillie, L. D. (2008). What is reinforcement sensitivity? Neuroscience paradigms for approach-avoidance process theories of personality. *European Journal of Personality*, *22*, 359–384. doi:10.1002/per.674

Smillie, L. D., Geaney, J., Wilt, J., Cooper, A. J., & Revelle, W. (2013). Aspects of extraversion are unrelated to pleasant affective reactivity: Further examination of the Affective Reactivity Hypothesis. *Journal of Research in Personality*, *47*, 580–587. doi:10.1016/j.jrp.2013.04.008

Spoont, M. R. (1992). Modulatory role of serotonin in neural information processing: Implications for human psychopathology. *Psychological Bulletin*, *112*, 330–350. doi:10.1037/a0024950

Stanton, K., Gruber, J., & Watson, D. (2017). Basic dimensions defining mania risk: A structural approach. *Psychological Assessment*, *29*, 304–319. doi:10.1037/pas0000337

Stein, D. J., Phillips, K. A., Bolton, D., Fulford, K. W. M., Sadler, J. Z., & Kendler, K. S. (2010). What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychological Medicine*, *40*, 1759–1765. doi:10.1017/S0033291709992261

Stepp, S. D., Yu, L., Miller, J. D., Hallquist, M. N., Trull, T. J., & Pilkonis, P. A. (2012). Integrating competing dimensional models of personality. *Personality Disorders: Theory, Research, and Treatment*, *3*, 107–126. doi:10.1037/a0025905

Stock, A.-K., & Beste, C. (2015). Conscientiousness increases efficiency of multicomponent behavior. *Scientific Reports*, *5*, 15731. doi:10.1038/srep15731

Suzuki, T., Samuel, D. B., Pahlen, S., & Krueger, R. F. (2015). DSM-5Alternative Personality Disorder Model traits as maladaptive extreme variants of the Five-Factor Model: An item-response theory analysis. *Journal of Abnormal Psychology*, *124*, 343–354. doi:10.1037/abn0000035

Tackett, J. L., Lahey, B. B., van Hulle, C., Waldman, I., Krueger, R. F., & Rathouz, P. J. (2013). Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *Journal of Abnormal Psychology*, *122*, 1142–1153. doi:10.1037/a0034151

Tackett, J. L., Quilty, L. C., Sellbom, M., Rector, N. A., & Bagby, R. M. (2008). Additional evidence for a quantitative hierarchical model of mood and anxiety disorders for DSM-V: The context of personality structure. *Journal of Abnormal Psychology*, *117*, 812. doi:10.1037/a0013795

Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Hashizume, H., Sassa, Y., … Kawashima, R. (2014a). Association between resting-state functional connectivity and empathizing/systemizing. *NeuroImage*, *99*, 312–322. doi:10.1016/j.neuroimage.2014.05.031

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., & Kawashima, R. (2014b). Regional gray matter volume is associated with empathizing and systemizing in young adults. *PLoS One*, *9*, e84782. doi:10.1371/journal.pone.0084782

Tang, T. Z., Derubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment. *Archives of General Psychiatry*, *66*, 1322–1330. doi:10.1001/archgenpsychiatry.2009.166

Turan, B., Guo, J., Boggiano, M. M., & Bedgood, D. (2014). Dominant, cold, avoidant, and lonely: Basal testosterone as a biological marker for an interpersonal style. *Journal of Research in Personality*, *50*, 84–89. doi:10.1016/j.jrp.2014.03.008

Valk, S., Bernhardt, B., Böckler, A., Trautwein, F. M., Kanske, P., & Singer, T. (2017). Socio-cognitive phenotypes differentially modulate large-scale structural covariance networks. *Cerebral Cortex*, 27, 1358-1368. doi:10.1093/cercor/bhv319

Wacker, J., & Smillie, L. D. (2015). Trait extraversion and dopamine function. *Social and Personality Psychology Compass*, *9*, 225–238. doi:10.1111/spc3.12175

Walton, K. E., Ormel, J., & Krueger, R. F. (2011). The dimensional nature of externalizing behaviors in adolescence: Evidence from a direct comparison of categorical, dimensional, and hybrid models. *Journal of Abnormal Child Psychology*, *39*, 553–561. doi:10.1007/s10802-010-9478-y

Wakefield, J. C. (1992a). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, *47*, 373–388. doi:10.1037/0003-066X.47.3.373

Wakefield, J. C. (1992b). Disorder as harmful dysfunction: A conceptual critique of *DSM-II-R's* definition of mental disorder. *Psychological Review*, *99*, 232–247. doi:10.1037/0033-295X.99.2.232

Wakefield, J. C. (1999). Mental disorder as a black box essentialist concept. *Journal of Abnormal Psychology*, *108*, 465–472. doi:10.1037/0021-843X.108.3.465

Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, *6*, 149–156.

Waller, N. G., DeYoung, C. G., & Bouchard, T. J. (2016). The recaptured scale technique: A method for testing the structural robustness of personality scales. *Multivariate Behavioral Research*, *51*, 433–445. doi:10.1080/00273171.2016.1157753

Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015). Extraversion and psychopathology: A facet-level analysis. *Journal of Abnormal Psychology*, *124*, 432–446. doi:10.1037/abn0000051

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist*, *62*, 71–83. doi:10.1037/0003-066X.62.2.71

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and statistical manual of mental disorders–fifth ed. *Journal of Abnormal Psychology*, *114*, 494–504. doi:10.1037/0021-843X.114.4.494

Wiener, N. (1961). *Cybernetics—or control and communication in the animal and the machine*, 2nd ed. New York, NY: MIT Press/Wiley. doi:10.1037/13140-000

Wouters, A. G. (2003). Four notions of biological function. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *34*, 633–668. doi:10.1016/j.shpsc.2003.09.006

Wright, A. G. C., Hopwood, C. J., Skodol, A. E., & Morey, L. C. (2016). Longitudinal validation of general and specific structural features of personality pathology. *Journal of Abnormal Psychology*, *125*, 1120–1134. doi:10.1037/abn0000165

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

Wright, A. G. C., & Simms, L. J. (2015). A metastructural model of mental disorders and pathological personality traits. *Psychological Medicine*, *45*, 2309–2319. doi:10.1017/S0033291715000252

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., … Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*, 1125–1165. doi:10.1152/jn.00338.2011

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