CHAPTER 52

PERSONALITY DISORDER AND VALIDITY: A HISTORY OF CONTROVERSY

PETER ZACHAR AND ROBERT F. KRUEGER

INTRODUCTION

There are very few, if any, constructs in psychiatry that are more conceptually encumbered than that of personality disorder. How so? Because personality disorder inherits not only all the problems associated with the concept of "personality," but also all the problems associated with the concept "psychiatric disorder." Personality alone encompasses biological-genetic, developmental, social-cultural, perceptual-cognitive, motivational-emotional, existential-phenomenological, and behavioral psychology. Most of the validity issues relevant to these areas are also relevant to personality and its disorders.

In this chapter we will review the historical development of the concepts of personality and personality disorder in psychology and psychiatry before turning to an exploration of several validity-related questions that intersect the interests of psychologists, psychiatrists, and philosophers. These include the role that values should play in the conceptualization of personality disorders, the nature of pathology in personality disorders, and the extent to which personality traits can or should be considered to be causal entities that carve nature at the joints.

FROM SELF TO PERSONALITY

Many modern definitions of personality emphasize that it is a pattern of organized psychological characteristics (e.g., traits, motivations, conflicts) that are stable across time and situations, that make individuals unique, and that determine thoughts, emotions, and behaviors.
Personality is the dynamic organization within the individual of those psychophysical systems that determine his characteristic behavior and thought. (Allport, 1961, p. 28)

Personality refers to an individual’s characteristic patterns of thought, emotion and behavior, together with the psychological mechanisms—hidden or not—behind those patterns. (Funder 1997, p.1)

One philosophically interesting claim in these definitions is that personality plays a causal role in the generation of thoughts, feelings, and behaviors. This claim predated the modern construct of personality, emerging in reflections on the nature of the self. Although a construct called the self began to appear in late antiquity, the psychological self as a bearer of thoughts and emotions was introduced only during the seventeenth and eighteenth centuries (Berrios and Marková 2003).

In nineteenth-century France, the term personality (personnalité) came to be seen by psychologists as a naturalistic alternative to metaphysical and quasi-spiritual notions about the unity and simplicity of the self (Lombardo and Poschi 2003). The academic philosopher turned psychologist Théodule Ribot believed the psychopathology of altered consciousness showed that self (or “personality”) was not a unity, and its components could be studied in a scientific way. At this time the concept of “dissociation” was introduced as a pathological counterpart to the empiricist’s notion that the association of ideas is the normal process by which minds are formed. The concept of dissociation was further developed by the neurologist Pierre Janet and later given significant credence in the USA by William James (1890) in his highly influential text The Principles of Psychology (Taylor 2000).

The Bostonian psychiatrist Morton Prince’s interest in what he called “multiple personality” was directly influenced by James’ appropriation of French ideas. Articles on personality appeared regularly in Journal of Abnormal Psychology which was founded by Prince in 1906 (Barenbaum and Winter 2003).

During the first decade of the twentieth century, character and personality were used interchangeably, but the term “character” carried moral connotations that came to be seen as inappropriate in a clinical context. One advantage of the term personality in mental health settings was that it allowed psychiatrists to put issues of morality and immorality aside in favor of non-judgmental clinical conceptualizations. After its adoption by the medical community in the USA, personality came to be considered to be more appropriate than character as a topic of scientific study—it was considered a natural psychological attribute like memory and perception (Danziger 1990; Nicholson 2003).

In 1921, Prince’s journal was renamed The Journal of Abnormal Psychology and Social Psychology and the psychologist Floyd Allport became its co-editor. The very first issue of the renamed journal included an article titled “Personality traits: Their classification and measurement” by Allport and his younger brother Gordon. Psychological measurement in the Galtonian tradition had been an alternative to experimental work as an approach to scientific psychology for at least a decade, but beginning in the 1920s it became an increasingly respectable alternative.

The development of intelligence tests for use in educational settings represented psychology's original foray into “psychometrics.” Danziger (1990) points out that it soon became clear that early intelligence tests had only a limited predictive value—even in educational settings. More importantly, many considered them less than helpful in the US military's...
efforts to assign draftees to positions in World War I; it would take additional efforts to arrive at modern, reliable and valid intelligence tests. After the war ended, it was suggested that personality traits might be more predictive of real-world behaviors than IQ scores. The assessment of personality traits was quickly adapted to the paper-and-pencil methods for assessing IQ; large groups that had been developed during the war and a new tool/instrument was born—the personality inventory.

The American Psychological Association accepted ownership of Prince's journal in 1926, leading to a large expansion of scientific articles in comparison to the clinical case reports favored by Prince (Barenbaum and Winter 2003). By the late 1930s the study of personality had become a unified content area in psychology, as signaled by the increased number of researchers who specialized in the study of personality, entire courses on personality being offered in university psychology departments, and the inclusion of a distinct chapter on personality in most introductory textbooks (Nicholson 2003).

This establishment of personality as a research specialty at this time is exemplified by the work of two men, namely, Gordon Allport (1937) and Henry Murray (1938). Many psychologists were involved in the development of personality as a specialization, but Allport and Murray, who both had the imprimatur of Harvard University, are often singled out.

Allport is associated with making the study of personality distinct from both clinical and social psychology. Although he followed the medical tradition by arguing for use of the secular term “personality” instead of “character,” he also advocated the study of the normal rather than the abnormal personality. His work in the measurement of normal traits represents the beginning of the empirical tradition in the psychology of personality. Until relatively recently, this tradition was segregated from psychiatric classification.

In addition to his concern with normality and health, Allport also emphasized the whole person and the subjective self. Psychologists later termed this foray into the study of normality humanism. Trait theorists, with their roots in faculty psychology, have always been interested in the biological basis of traits. In contrast, the humanistic psychologists were primarily interested in exploring meaning and personal values. In the 1950s, humanistic psychology was embraced by the rapidly expanding specialties of clinical and counseling psychology (Rogers 1951). Like the academic personality researchers, the humanistic clinical psychologists also had limited interest in psychiatric classification and diagnosis.

Unlike Allport, Murray was not trained as a scientific psychologist and he was a hostile critic of conventional academic psychology (Nicholson 2003). Nor did his preference for case studies and case conferences endear him to the Harvard psychology department. Yet, Murray's Freud-inspired theory of needs became influential in clinical psychology, as did other psychodynamic models including those of Jung, Adler, Horney, and Fromm (Hall and Lindsey 1957). What became of this tradition will be briefly explored at the end of the next section.

To summarize this section, the construct of personality represents a secularization of “character” for use in a medical context. It is also an heir to philosophical theories of the subsistent self. As a research specialty in psychology, personality was “normalized” and considered to be qualitatively distinct from psychopathology. Two forms this normalization took was the psychometric measurement of traits and the humanistic concern for individual subjectivity and meaning.
THE HISTORY AND THE CONCEPT OF DISORDER

The history of personality disorder begins with the origin of modern psychiatry in the nineteenth century. Our story commences with the French physician Benedict Morel's introduction of dementia praecox in the 1850s, which he tied to the notion of degeneration. According to Morel, degeneration was an irreversible physical and mental deterioration from a higher to a lower form. Morel also believed that degeneration represented a fall from grace in the theological sense (Gilman 1985) and referred to the signs of degeneration as *stigmata* (Carlson 1985). Morel believed that once the degenerative process got going in a family, the offspring became increasingly degenerate. Those initially affected were nervous and violent. Subsequent cohorts were more likely to be epileptic, hysterical, and hypochondriacal. As degeneration proceeded, family members became eccentric and unpredictable, with dementia praecox being one of the later, final manifestations of the decline (Carlson 1985).

It is common to consider dementia praecox (later renamed schizophrenia) and its associated psychotic symptoms (especially delusions) as the prototype of insanity. It is, however, also historically accurate to say that the hypothetical process of degeneration itself was for a time an equally important theoretical model of insanity, including within its scope the notion of the morbid personality.

As degeneration theory spread it came to refer to deviates of all stripes and colors, including sexual deviates, criminals, the mentally retarded, and so-called inferior races. The moral dimensions to these categories were readily identified as character flaws. Prominent figures in the study of "degenerates" in the 1870s include Valentin Magnan in France, Cesare Lombroso in Italy, Henry Maudsley in England, and Richard von Krafft-Ebing in Austria (Carlson 1985; Pick 1989; Shorter 1997, ). Depending on the thinker, the presumed etiological factors included alcohol and drug use, masturbation, poisoning, harsh climates, harsh working conditions, poor diet, neglect, and parental turmoil.

Crucially important was the integration of the degeneration hypothesis with mid-nineteenth-century evolutionary theory, whereby degeneration became a natural rather than a theological concept. This process was facilitated by the 1844 publication of *Vestiges of the Natural History of Creation*, authored anonymously by Robert Chambers. It was Chambers, not Darwin, who introduced the larger British public to evolution and the naturalistic approach to history. Referring to the *Vestiges* as a "Victorian sensation," Secord (2000) reports that the book was widely read (and condemned) by all strata of society and the public scandal led Darwin to delay publication of his own evolutionary theory for fourteen years.

In the *Vestiges*, Chambers claimed that evolution could be both progressive and regressive (i.e., degenerate). This bidirectional notion can be traced back to the origin of evolutionary thinking in the speculative work of Buffon in the eighteenth century, but through *Vestiges*

---

1 Carlson (1985) notes that the French thinker Abbé Raynal included emigrated white Americans in his examples of degeneration. Both Benjamin Franklin and Thomas Jefferson objected.
the degeneration hypotheses permeated British Society. For example, in the literature of this time Mr Hyde, Count Dracula, the half-human half-animal creatures of Dr Moreau, and the Morlocks from The Time Machine were all depictions of degenerates (Pick 1989).

An additional nineteenth-century development important to the construct of personality disorder was the extension of the term “insanity” to cases that did not involve what currently might be called a disorganized psychosis. For example, in 1801 Pinel introduced the classification manie sans délire. Berrios (1993) notes that the meaning of these terms have changed so much that we cannot assimilate them to our current notions of mania and delusions, and that Pinel’s new classification primarily referred to a deviation from a state of “total insanity” in which rationality was somewhat preserved. Depression without psychotic features and obsessions and compulsions would be included in this category. Esquirol’s 1810 notion of monomania was an alternative conception of partial insanity. Monomania explained some kinds of abnormal behavior by positing a single, fixed delusion that had radiating effects throughout consciousness and behavior, but did not entail a comprehensive disorganization of mental faculties (Goldstein 1987). J. C. Prichard’s 1835 concept of moral insanity, which Berrios (1993) claims bears no resemblance to the modern construct of psychopathy/antisocial personal disorder, also belongs in this tradition of classifying non-psychotic mental disorders.

The notions of degeneration, character, and partial insanity came together most explicitly in J. A. Koch’s 1891 concept of psychopathic inferiority. Koch used personality traits such as trouble-maker, arrogant, morose, and touchy to describe these conditions, the more severe forms of which he considered to be the result of a process of degeneration (Berrios 1996; Schneider 1923/1950).

The work of Emil Kraepelin was a major influence on many aspects of psychiatric thinking both before and after the turn of the century. In Lectures on Clinical Psychiatry, Kraepelin (1904) described the morbid personality as a state of permanent psychopathic inferiority—a term he used for stable and unremitting conditions with a hereditary basis. Morbid (or “degenerate”) personality specifically referred to a group of narcissistic, irresponsible, and disagreeable individuals:

If such a personality is measured by the standard of a law-court, it is simply that of a criminal and a swindler. Yet the physician cannot escape from the conviction that the patient has a congenital incapacity for a regular course of life, stronger than all education, experience, and self-control (Kraepelin 1904, p. 302).

As Kraepelin notes, such patients would seem normal in a casual encounters or upon a first meeting. The morbidity could only be gleaned by looking at the pattern of a whole life.

At some point in the first decade of the twentieth century, the tide against degeneration began to turn in psychiatry. This was likely a combination of improved information about hereditary processes in psychiatry and elsewhere, and growing philosophical disagreements with the social and political claims of its most committed advocates. Kraepelin largely abandoned the hypothesis by 1918 (Carlson 1985).

By the time that Kurt Schneider wrote Psychopathic Personalities in 1923, degeneration theory, and with it the notion of morbid personality, were much less tenable. For Schneider (1923/1950), “psychopathic personality” was a general term for all personality disorders. He considered these to be clinical, not moral conditions. Schneider defined abnormal personality in an objective, statistical sense as deviation from an average range. What made an
abnormal personality clinically-relevant (psychopathic) was whether it led to either personal suffering on the part of the patient or to the suffering of others. Schneider also stated that psychopathic personalities were conditions that lay outside the narrow confines the disease-oriented approach to medicine and were not literal illnesses.

One additional contribution to the historical development of the construct of personality disorder must be mentioned. Late in his career, Freud (1921/1960) proposed what has come to be known as his structural (id, ego, and superego) model of personality. The development of this theory by thinkers such as Reich (1949) and Hartmann (1958) became influential in the USA. In particular, clinical psychologists who specialized in diagnostic testing came to see their task as developing a model of the patient's personality (Reapapart et al. 1945; Shapiro, 1965). This applied not only to hysteria and paranoia, but also to depression and schizophrenia. The third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III; American Psychiatric Association 1980) in 1980 is often given credit for raising interest in personality disorders by separating their diagnosis from the diagnosis of other mental disorders, but it is rarely noted that in doing so the DSM-III created a greater distinction between personality and other mental disorders than previously existed in the American approach to psychopathology.

**PHILOSOPHICAL ISSUES AND VALIDITY**

The philosophical issues relevant to the validity of personality disorders can be parsed in different and partially overlapping ways. Many of the issues discussed in the psychological and psychiatric literature do not map neatly onto the professional concerns of philosophers. For example, the technical issue of whether a personality construct is distributed in the population as a true class (a "classical category") or as a continuous dimension dominates the scientific literature, but barely registers on the philosopher's own conceptual radar.

The three validity issues we will explore are rooted in the history just reviewed and they also represent ongoing concerns in psychology and psychiatry. They were selected because they could benefit from sustained philosophical exploration, i.e., explicit philosophical work has an important role to play in the development of solutions/resolutions. Furthermore, these issues are readily registered on the radar of psychiatrists, psychologists, and philosophers.

**Issue one: What role do values and considerations of character play in the conceptualization of personality disorders?**

Consider this passage from Jonathan Franzen’s novel *Freedom*:

> A game could be made of trying to get Patty to agree that somebody's behavior was “bad.” When she was told that Seth and Merrie Paulsen were throwing a big Halloween party for their

---

2 For example, "Axis I" would code for psychiatric disorders such as depression. "Axis II" would code for personality disorders and personality traits. "Axis III" would code for medical conditions such as type II diabetes.
twins and had deliberately invited every child on the block except Connie Monaghan, Patty would only say that this was very "weird." The next time she saw the Paulsens in the street, they explained that they had tried all summer to get Connie Monaghan's mother, Carol, to stop flicking cigarette butts from her bedroom window into their twin's little wading pool. "That is really weird," Patty agreed, shaking her head, "but you know it's not Connie's fault." The Paulsens, however, refused to be satisfied with "weird." They wanted sociopathic, they wanted passive-aggressive, they wanted bad. (Franzen 2010, pp. 5–6)

As will be reviewed in more detail later in this chapter, many psychologists believe that the universal structure of human personality is best modeled in terms of five broad traits—called by some The Big Five. Interestingly, when psychologists attempt to model the structure of normal personality at levels higher than the Big Five, they find two superfactors that Digman (1997) terms Alpha (α) and Beta (β) and De Raad and Barel (2008) call virtue and dynamism. Virtue is said to capture the moral aspects of character such as kindness, reliability, and decency. Dynamism encompasses features such as enthusiasm and energy. Such findings suggest that despite the historical attempt to naturalize the morally-laden notion of character by introducing the construct of personality, the moral aspects have not been eliminated.

Much the same is true in the realm of psychopathology. DeYoung (2010) has termed Alpha "stability" and Beta "plasticity," and has noted that the best markers of Alpha are descriptors with negative valence (e.g., "I get out of control"), whereas positively valenced descriptors correspond with Beta (e.g., "I express myself easily"). Simms, Yufik, and Gros (2010) have shown empirically that negative valence is a potent predictor of personality disorder, whereas positive valence is considerably less relevant. Essentially, personality disorder as conceptualized in the contemporary psychiatric literature pertains to the tendency to behave in a negatively valenced fashion. Character and virtue are therefore inextricably interwoven into the fabric of personality disorders.

It is in this light that one should consider a variety of positions on the "mad versus bad" problem with respect to personality disorders. John Sadler (2005) notes that modern societies have to be careful about medicalizing what they consider immoral behavior. He is particularly concerned about those conditions that most people would consider to represent violations of moral norms. For example, every single criterion for identifying conduct disorder such as cruelty and theft represents a violation of moral norms. This does not mean that conduct disorder is not a legitimate disorder, but the possibility that bad behavior has been inappropriately medicalized should not be dismissed.

In the domain of personality disorder, Louis Charland (2004, 2006) has argued that the Cluster B personality disorders such as borderline, narcissistic, and antisocial are morally bad groupings of behaviors that have been inappropriately medicalized. Charland argues that if all the morally-laden diagnostic criteria for these conditions were eliminated, there would be nothing left to diagnose. Furthermore, successful treatment of these "disorders," he claims, would be more akin to moral conversion.

Such assertions have to be considered with some care. As we have seen, after the abandonment of the construct of the morbid personality, personality disorders have not been considered to be disease-like entities along the lines of schizophrenia. Their "medical" status resides on the border between the dysfunctional and the disliked. Furthermore, if any

3 Excerpt from "Good Neighbors" from Freedom by Jonathan Franzen. Copyright © 2010 by Jonathan Franzen. Reprinted by permission of Farrar, Straus and Giroux, LLC.
particular personality pattern is considered to lie within the fuzzy boundaries of the medical (i.e., be dysfunctional), there is no reason for thinking that the relevant capacity failures cannot include moral capacities.

Zachar and Potter (2010) respond to Charland by claiming that, instead of seeing moral theory as a threat to the study and treatment of personality disorders, moral theory can appropriately inform psychiatric thinking about personality disorders. In particular they explore the conceptual resources of virtue theory.

According to Zachar and Potter the concept of flourishing in virtue theory has a much larger scope than those philosophical notions of “morality” that emphasize prescribing law-like rules for governing behavior. Virtue theorists, who emphasize the importance of thinking about what kind of persons we should be rather than thinking about what rules should govern our behavior, would not expect that notions of morality, goodness, health, and adaptivity can be completed segregated.

Zachar and Potter also argue that if one decides to conceptualize traits along the lines of virtue theory with pathology/vice existing at the poles and virtue/health being in between, one should not be too concrete about the evaluative implications of any position within that structure. The golden mean is not an arithmetic mean but a dynamic balance point that represents a “better” (not perfect) response given all one’s competing obligations and goals. The adaptiveness of any trait should be evaluated in the context of the person’s situation. It should also be evaluated in the context of other traits which can compensate for or exacerbate any negative impact on adaptiveness.

One does not have to agree with Charland regarding the “moral, not medical” nature of personality disorders to be concerned about medicalization. Carl Elliott (1996, 2003), for example, withholds final judgment about the pathological status of personality and related states of character, but remains concerned about the consequences of medicalizing common personality traits. Part of his concern, very interestingly, relates to people’s ambivalent attitudes about medicalization. Contemporary society cannot seem to decide whether medicalization (e.g., conceptualizing debilitating shyness as social phobia) is a scientifically improved or a shallower way to think about traits.

The risk of over medicalization is greater when normal personality traits are included in the assessment and the scales are bipolar. Consider what has arguably become the dominant test for measuring normal personality, the NEO-Personality Inventory (NEO-PI-R and recently NEO-PI-3). Some NEO-PI-R advocates conceptualize traits as being structured like virtues, with the extreme ends of each dimension being considered pathological (Widiger and Mullins-Swatt 2009). The problem with most bipolar personality dimensions such as those in NEO-PI-R is that they were not designed specifically to measure psychopathology (Krueger et al. 2011). For example, the NEO-PI-R has scales to measure varieties of “agreeableness” but not scales to specifically index the conceptual opposite of agreeableness, i.e., frankly hurtful and antagonistic interpersonal behavior.

One of the flaws of many early clinical tests, particularly projective instruments such as the Rorschach Inkblot Test, was that nearly everyone given the test in a clinical situation was judged to be pathological because the examiners were looking for psychopathology. This overpathologizing problem potentially applies to using the NEO-PI in a clinical context as well. For example, consider the NEO Problems in Living Checklist which is sold with the third edition of the test (NEO-PI-3). The Checklist provides pathological interpretations for the extreme end of every scale, including all thirty facet scales (i.e., the scales have a bipolar structure).
structure). An example is the agreeableness facet named “trust.” High scores can potentially signal being gullible and naïve and low scores signal being paranoid and jealous. With bipolar dimensional models there is also a risk that moderately high scores can be dubbed “sub-threshold,” and be considered potentially pathological. The authors of the test, Costa and McCrae, are also concerned about over-medicalization and caution again using the Checklist in this “cookbook” way, noting that these are only possibilities that must be independently confirmed (Costa and McCrae 2010). A more direct way to attenuate inappropriate medicalization is to calibrate trait scales to measure psychopathology itself. This is the approach taken by tests such as the Minnesota Multiphasic Personality Inventory (MMPI).

Another import feature about the latest edition of the MMPI that addresses some of the concerns of Charland and Elliot is that the scales are unipolar not bipolar. When scales are unipolar, high scores can be considered pathological whereas low scores generally have no specific interpretation. Many people can take the MMPI and have no interpretable scores, i.e., they are not experiencing the kinds of clinical problems that are assessed by the test. It is an empirical question as to whether high and low scores can be interpreted, but the advantages of designing scales to be unipolar should not be dismissed. This does not completely solve the problem of over-medicalization. Unipolar scales interpreted in a context-free manner can be harmful as well, but relative to bipolar scales that assess normal personality, the risk of harm is reduced.

**Issue two: What makes personality “pathological”**

One of the problems that arose after personality disorder was separated from degeneration theory and the disease model was justifying its pathological status. As might be expected, a plurality of models are used to justify the attribution of “disorder” to personality. Some of these models require minimal inferences and tend to be less controversial. The more inferences required, the more debatable the model. Those models that view personality disorders as inherently pathological rather than clinically-relevant are the most debatable. We will review six models. Our listing of the models is based on a discussion by Zachar (2011) which itself was influenced by conceptual work on the relationship between personality disorders and other psychiatric disorders (Dolan-Sewell et al. 2001), and on the relationship between temperament, personality traits, and personality disorders (Clark 2005).

The **vulnerability model** claims that personality disorders are clinically-relevant conditions in the same way that conditions such as hypertension and hypercholesterolemia are clinically relevant. They are clinically relevant because they are risk factors for the development of less controversial disorders. For hypertension the increased risk is for cardiovascular and cerebrovascular disease. For personality disorders the increased risk includes depressive disorders, anxiety disorders, eating disorders, and psychosis (McGlashan et al. 2000; Oldham et al. 1995).

The **patho-plantarity model** claims that personality disorders are clinically relevant conditions because they affect the course and outcome of other psychiatric disorders. Not only are people with personality disorders more vulnerable to developing other psychiatric disorders, when they do develop them they tend to develop them earlier in life, have more severe symptoms, and worse outcome (Clark 2007).

The **spectrum model** claims that personality disorders represent milder expressions of the same genetic predispositions that underlie more serious disorders (Lenzenweger 2006;
Meehl 1962). An early version of this model favored by Kraepelin (1907) and Kretschmer (1925) held that personality disorders represent milder manifestations (formes frustes) of serious psychiatric disorders. Examples of personality disorders that have been hypothesized to exist on a spectrum include schizotypal, schizoid, and paranoid personality disorders (schizophrenic spectrum) cyclothymic personality disorder (bipolar spectrum) and depressive and anxious personality disorders (internalizing spectrum) (Kendler et al. 1993; Phillips et al. 1995; van Valkenburg et al. 2006).

The decline in functioning model is best represented by degeneration theory, i.e., personality disorder represents a developmental unexpected decline in functioning (or morbid change). As we have seen, the very idea of regression to a more primitive type (atavism) is no longer considered an empirical possibility. A residue of degeneration may be detected, however, in the concept of genetic anticipation—which occurs in trinucleotide repeat disorders such as Huntington’s disease (Carpenter 1994). The more repeats, the earlier the onset of the disease, and in theory the number of repeats can increase over generational time. The risk that genetic anticipation could result in some eugenics-like social policies is always a possibility, although at this point finding empirical support justifying its application to personality disorders seems far-fetched.

Another possible formulation of the decline in functioning model would be personality change due to: (a) severe emotional trauma, (b) a previous psychiatric illness, or (c) a general medical condition such as a brain injury. In these cases the personality-relevant symptoms are the same kinds of symptoms that define DSM and ICD personality disorders. All these overlapping symptom clusters are potentially maladaptive, with a main difference being that some clusters of symptoms (such as seen in brain injury patients) can be explained with respect to an aberrant causal history where other clusters (the personality disorders diagnosed in the DSM) cannot. What counts as an aberrant causal history, however, is a vexing philosophical problem. Does sexual abuse count? If so why do not all victims of sexual abuse develop personality disorders? As George Graham (2010) might say—experiences of abuse do not always “gum up the works”—one has to consider not only causal events, but also the particular pattern of vulnerability and resiliency variables on which they act.

The impairment-distress model is favored by Kurt Schneider (1923/1950), proponents of the five-factor model (Widiger 2006), and the authors of the DSM-IV. In addition to statistical abnormality, advocates of this perspective argue that personality styles and traits can be considered disordered if they reliably lead to distress or impairment in social and occupational functioning. In some versions of this model, the impairment is partly explained with respect to inflexibility/dyscontrol (Widiger 2006).

The capacity failure model emphasizes the failure to develop one or more psychological capacities that contribute to normal functioning. Unlike the decline in functioning model where capacities are lost, according to this model they never develop. It is therefore expected that the problematic symptoms associated with a personality disorder will be evident from a young age, often by the teenage years. Personality disorders in this sense are not like acquired conditions such as tuberculosis or major depression. One could think of them as developmental disorders similar to autism, except that they tend to emerge in adolescence rather than childhood. This is more explicit in the DSM-IV, with conduct disorder considered as necessary to diagnose antisocial personality disorder, but it is presumed for all other personality disorders.
As is true with all models that specify the nature of the dysfunction, there must be some notion of what counts as normal functioning—which requires speculation. Two popular ways of conceptualizing "normal" are the natural function approach (Millikan 1984; Wakefield 1992) and the causal-role approach (Cummins 1983). Natural functions refer to functions that exist because they were produced by natural selection. The causal-role approach is ahistorical and more mechanistic. Function refers to the causal contribution that a specific capacity plays in the overall functioning of an organism.

Some recent research has begun to make some headway on the problem of modeling capacity failure. For example Roel Verheul and colleagues note that a high level of neuroticism is not itself a capacity failure. In order to move toward a more substantial notion of personality pathology, they present a list of sixteen normal capacities that are relevant with respect to personality disorders (Verheul et al. 2008). These capacities include emotional regulation, purposefulness, enduring relationships, responsible industry, and cooperation. They further propose that severe personality disorders are cases in which a large number of these fundamental capacities are compromised.

Based on this brief survey, we conclude that a plurality of models is needed for understanding the nature of personality pathology. No single model seems able to adequately justify the pathological status of personality disorders, and those that attempt to do so more directly require making theoretical inferences that go beyond the evidence. More work on an adequate theory of pathology for personality is still required.

Issue three: Is personality a causal entity?

One aim of our selective historical review was to contextualize this third, very difficult issue. Psychologists who study personality from the standpoint of trait theory claim that the contention of Allport that normal and abnormal are qualitatively distinct has been empirically refuted—not only with respect to personality, but with respect other disorders such as depression and schizophrenia (Clark 2005; Livesley et al. 1998; Smith and Combs 2010). For example, the current best evidence indicates that the distinction between depressed mood and normal mood does not represent a carving of nature’s joints, but a practically useful cut on an underlying continuum of mood similar to a distinction between short and tall (Watson 2005). Those who emphasize what psychologists call the quantitative (difference in degree) rather than the qualitative (difference in kind) distinction between normal and abnormal are termed dimensional model proponents.

Psychologists who advocate a dimensional approach are also inching toward making personality foundational to the conceptualization of psychopathology in general as did the Freudian ego psychologists mentioned previously, but they remain skeptical about the medical model approach of the psychodynamic tradition which emphasized case studies and the identification of personality types such as borderline, histrionic, and narcissistic. The new personality-oriented psychopathologists argue that the symptom space of personality pathology can be more comprehensively described by empirically-supported trait models (Trull and Durrett 2005; Widiger and Trull 2007).

Let us introduce the issue of the casual role of traits by briefly turning to philosophical history. In his *A Treatise of Human Nature* the empiricist philosopher David Hume argued that we have no experience of the existence of a self. All we have experiences of are organized
bundles of sensations and perceptions. These bundles succeed each other in a systematic way and are more-or-less similar from one to the next (in our memory). What we call the self, says Hume, represents the reification of that similarity into a subsistent entity, but it is not a real causal entity.

Behavioral observations indicate that people who are anxious in many situations are also likely to be fearful, self-conscious, and emotionally overwhelmed—supporting an inference that they are high on the dimensional construct of neuroticism. According to an empiricist, the dispositional term neuroticism encodes information about probabilities. When it is doing its job, neuroticism is an inference ticket for knowing which behaviors are correlated with each other and for developing expectations about what patterns are more likely to occur in the future. For empiricists, the hypothetical construct of neuroticism is an instrument that we use to organize and navigate information about human behavior, but it is not a real entity in the head any more than the subsistent self was an entity in the head for Hume.

As we say in early this chapter, the conceptual heirs of the subsistent self in modern scientific psychology are personality and personality traits. In the science of psychometrics, personality traits are considered to be latent variables. Latent variables are mathematical abstractions that represent hidden causal structures that determine motives, perceptions, interests, and behaviors. Latent variable models lead contemporary trait theorists, in contrast to the instrumentalism of the empiricists, to adopt scientific realism about personality traits. As we will see in the following sections, trait realists believe that universal causal models for psychology are likely to be grounded in biology. From their perspective neurotic traits such as being fearful and self-conscious are only surface properties. Neuroticism proper, they believe, refers to an endogenous physiospsychological entity that exerts a causal influence—it produces the neurotic pattern of behavior.⁴

---

**The received view: Traits as causal entities in the person (the FFM perspective)**

What Whitehead (1926) named the fallacy of the perfect dictionary is the hypothesis that humans have considered all the fundamental ideas that are relevant to understanding experience and encoded them in language. The perfect dictionary hypothesis, however, remains the underlying assumption of the lexical approach to trait theory. For example, the lexical theorist Raymond Cattell (1943b, 1943a) claimed to have developed a representative sample of the population of personality traits, which he then submitted to statistical analysis to see if it could be reduced to a small number of general constructs with minimal loss of information. Cattell’s (1945) original model had twelve dimensions.⁵ Advocates of the received view claim that it slowly became clear that trait models containing numerous and specific personality variants could be collapsed into a more general five-dimension model (Fiske 1949;

---

⁴ In contrast to the philosophical empiricist’s emphasis on the particular, the contingent, and the probable, most contemporary trait theorists are seeking universal, necessary, and relatively certain knowledge of why things happen. Hence, they could be considered rationalists rather than empiricists.

⁵ The twelve factors were derived from data gathered by observer ratings. When the factors were measured by test items and the data analyzed was based on self-report, four additional factors for a total of sixteen emerged.
Norman 1964; Tuples and Christal 1958). Different studies they claim, independently discovered the same five dimensions, irrespective of whether the data were based on self-report or peer observations. Termed the “Big Five,” these dimensions are called: (1) surgency (extroversion), (2) agreeableness, (3) conscientiousness, (4) neuroticism, and (5) intellect/openness to experience.

In the late 1980s, Paul Costa and Robert McCrae developed a personality test to measure the Big Five (called the NEO Personality Inventory, NEO-PI). Rather than seeking to discover personality factors, they set out in an a priori fashion to measure what had already been identified by the “inductive” methods of the lexical theorists. In adopting an explicitly scientific realist position regarding the nature of traits, Costa and McCrae have become the best-known representatives of the received view. This is exemplified in their introduction of the term “five-factor model” (FFM)—by which they mean a comprehensive scientific (explanatory) model of personality (McCrae and Costa 2008).

According to the FFM narrative, subsequent to the model’s introduction, these traits were shown to be highly stable over a period of several years (McCrae and Costa 2003). FFM advocates also report that the model has been replicated across all cultures studied so far, suggesting that the traits are domain general personality factors (McCrae and Costa 2008). The stability and cross-cultural consistency of the traits has led to a judgment of universality (McCrae and Costa 1997). The additional discovery that personality traits have genetic bases has resulted in the assertion that these traits are grounded in nature not culture or experience (McCrae 2004).

A plurality of contrarian views

Independent discovery?

What responses do contrarians have with respect to the received view? One response is to dispute the FFM discovery narrative. In this narrative, the five-factor model is psychology’s version of the discovery of the J/ψ particle in physics: different research teams working independently with separate methodological strategies made the same discovery. As happened in physics, this narrative is taken to justify adopting a realist view about the discovered entities.

In contrast, Block (1995) argues that rather than starting from scratch and finding the fixed world structure waiting to be discovered, the various psychological research programs ended up in the same place, roughly, because they began in the same place, roughly. They were looking for a five-factor model based on clues from a shared set of previous findings. In designing the research, trait terms were also included and excluded with such an outcome in mind. What happened, according to Block, was not independent discovery, but planned replication.

Universality?

Here are McCrae and John (1989) addressing the question of why there are five factors:

We believe it is simply an empirical fact, like the fact that there are seven continents on earth or eight American presidents from Virginia. Biologists recognize eight classes of vertebrates (mammals, birds, reptiles, amphibians, and four classes of fishes, one extinct), and the theory of evolution helps to explain the development of these classes. It does not, however, explain
why *eight* classes evolved, rather than four or eleven, and no one considers this a defect in the
theory. There are, of course, reasons why human beings differ along each of the five personality
dimensions—reasons to be found somewhere in evolution, neurobiology, socialization, or the
existential human condition. But it is probably not meaningful or profitable to ask why there
happen to be just five such dimensions. (p. 194)

A contrarian can dispute such scientific realism by questioning the universality assumption. Several thinkers point out that although the five-factor model has been found in other
cultures, so have alternative models, including three-, six-, seven-, and eight-factor models
(Almagor et al. 1995; Ashton and Lee 2001; Clark and Watson 1999; De Raad 2009; De Raad
and Barelds 2008). Advocates of models with greater than five factors argue that meaningful
aspects of personality (such as honesty and hedonism) are distinct factors (Ashton and Lee
2001). Along a similar line of reasoning, it has been argued that the FFM is blind to the trait
of schizotypy/psychoticism (Harkness et al. 1995; Shedler and Westen 2004; Zachar 2008).

De Raad and Barelds note (2008) that different combinations of personality items tend
to emerge depending whether the data being analyzed are verbs, nouns, trait adjectives, or
some combination of all three. Almagor et al. (1995) also state that different dimensions
emerge depending on whether evaluative terms are included or excluded. Finally, Simms et
al. (2010) provide evidence that these evaluative dimensions contribute additional informa-
tion, beyond the FFM, that is relevant to the prediction of personality disorder.

With respect to the current chapter the most important example of this complex relationship
between the kind of data used (input) and the solution (output) is what happens when the
data are psychiatric symptoms. One of the most consistent findings in the past fifteen
years is that patterns of comorbidity among many mental disorders can be accounted for by
two superordinate dimensions, internalizing (INT) and externalizing (EXT). Comorbidity
refers to the tendency for putatively distinct mental disorders to co-occur more often than
predicted by chance. INT and EXT were originally derived in the work of the psychologist
Thomas Achenbach on the structure of psychiatric problems in children (1995) and have
turned out to be fruitful in understanding disorder co-occurrence in adults (Krueger 1999;
Krueger et al. 1998). Put simply, internalizing disorders involve experiencing emotional dys-
regulation—it is what depression and anxiety disorders have in common. Externalizing disor-
ders involve acting out emotional dysregulation in ways that most societies deem illegal.
It is what antisocial personality disorder (doing things that would be grounds for arrest) and
illicit substance dependence have in common.

Eaton et al. (2010) argue that INT and EXT are stable, present across cultures, and genet-
ically meaningful. These dimensions are not, however, derived from the FFM in a direct
fashion. This is because the FFM is derived from everyday language for describing normal
personality, whereas INT and EXT are derived from mental disorder definitions used by psy-
chiatrists. From the FFM standpoint, INT and EXT are likely blends of personality factors
and therefore heterogeneous—for example, both are correlated with neuroticism (Krueger
and Markon 2006). Externalizing also shares variance with antagonism (low agreeableness)
and negligence (low conscientiousness (Krueger et al. 2011)). Within the domain from which
they were derived (mental disorders), however, INT and EXT are single and coherent fac-
tors. Returning to the geographical analogy, this would be like discovering seven continents
while surveying them on a ship, but only two while surveying them from an airplane.

One of the goals of the scientifically oriented psychopathologists is to propose explicit
mathematical models of various “symptom spaces” and compare them with each other in

*Biological* Another interesting question to ask is whether this research can be traced to biological, temperamental, or environmental differences.

A large number of genetic factors (Krugier et al. 2000) influence mental disorders. It is
noteworthy that the common twin studies have consistently found that genetics is a major
influence in which identical twin pairs have the same mental illness.

For example, a genetic factor

---

6 It is worth noting that the factor models, as partial
terms of adequacy. One could call this a *neo-Galilean* approach—meaning the goal is to construct mathematical models that best account for the subject matter being studied. Models that have what Krueger calls *structural validity* (Krueger and Eaton 2010) should fit the data according to specific mathematical criteria of adequacy (termed “goodness of fit”). Although psychiatry has typically used a natural history approach along the lines of Sydenham and Darwin, these recent psychometric developments in mathematical modeling constitute a research tradition that is more similar to modeling approaches used in other sciences.

Interestingly, the Big Five tradition is more mathematically-based than is the FFM for two reasons. First, the “FFM” is typically not associated with a model that transcends operationalization, but rather, with a single instrument—the NEO-PI-R (Costa and McCrae 1992). Abstract models and their concrete measures are not the same thing. Models are partial representations of particular domains⁶ (Giere, 1999), but as we have noted, some thinkers have interpreted the NEO-PI-R scales as being identical with the five factors. Second, an important feature the NEO-PI-R is its division of the five factors into subscales (facets). For example, anxiety, angry hostility, and depression are facets of the neuroticism factor. This hierarchy, however, was derived rationally, as opposed to being inductively found in the data as were the Big Five. It may even be that clinically useful facet scales are too specific to emerge in models that capitalize on partialing shared variance (Krueger et al. 2011). This makes facets more like countries than continents.

To conclude this section, while the successful replication of the five factor model across different samples is impressive, there is too much variability in various factor structures to say that there is one and only one universal and fixed structure of personality consisting of extraversion, conscientiousness, agreeableness, neuroticism, and openness. In scientific realist lingo the five-factor model is at best an approximation to a circumscribed reality.

**Biological, therefore, carving nature at its unique joints?**

Another important assumption of the received view is that if a trait can be shown to have a biological basis, then it is both scientifically valid and “real.” A crucial area of research in this respect has been behavioral genetics. The workhorse of behavior genetics has been the twin study, which involves comparing identical and fraternal twins to get a handle on genetic and environmental contributions to human individual differences (Plomin et al. 2008).

A large number of reliable personality constructs seem to have heritabilities around 0.50 (Krueger and Eaton 2010), meaning that half of the total variation in all personality traits can be traced to genetic differences among persons. For example, basic aspects of emotional temperament such as neuroticism are heritable to essentially the same degree as individual differences in the tendency to resonate with art and literature (i.e., openness to experience).

Why there seems to be so little variance regarding the heritability of personality traits raises interesting methodological and conceptual questions. One possibility, pursued by Krueger et al. (2008) is that the 50% figure is an average, across a range of “micro-environments” in which the heritability might differ notably. More technically, genetic and environmental influences on a phenotype such as personality are moderated (changed by) other variables.

For example, Krueger et al. (2008) examined the extent to which genetic and environmental influences on personality vary as a function of the relationships between adolescents

---

⁶ It is worth noting that mathematical models are partial representations of theories just as theories are partial representations of some substantive domain.
(seventeen-year-old twins) and their parents. They found that, averaging across parent-child relationships, the heritability of neuroticism was 50%, similar to other findings in the literature. However, this 50% figure averaged across a wide range of relational scenarios that had a notable impact on the genetic and environmental architecture of personality variation. In relationships where there was a high level of conflict, the heritability of neuroticism was considerably diminished, and the level of family environmental influence was as strong as the level of genetic influence. This suggests that "the environment" and "genes" may not have consistent direct effects on personality, but rather, that these forces are in dynamic interplay. In environments where family conflict existed, that environment was relevant to personality variation. One therefore has to be careful about generalizing heritability estimates from one sample to the next, and also about assuming that a high heritability estimate proves that the trait in question is carving nature at its systematic joints.

Latent variables as causal entities?

Other critics question the essentialist assumptions underlying the latent trait model (Cramer et al. 2010; Kendler et al. 2011). Analogous to the notion of a single factor called pure g that underlies all individual differences in academic ability, trait essentialists assume the existence of a quality of pure neuroticism which some people display more of than others. In McCrae and Costa’s model, how much neuroticism someone has is fixed by biology.

According to psychometric theory, test items are fallible indicators that can triangulate on the latent trait. That is to say, when psychologists measure neuroticism using a scale composed of many items (e.g., life is a strain for me), everyone who scores very high or very low on the scale could be said to have or lack this quality. In theory, different items can possess higher or lower doses of the latent trait of neuroticism and should be worth more or less points. People who are average on the trait should get there by agreeing with the average and low dose items, and not agreeing with high dose items. There is also promising evidence that constructing scales of this type is an empirical possibility (Watson et al. 2007, 2008).

Following up on a point made by Grove and Vrieze (2010), however, one must also consider the possibility that as long as items are preselected to be unidimensional, such a "dose response" structure is itself preselected. If a theory is used to guide the selection of empirical data, then to what extent can the data be said to confirm the theory? This is a difficult question. At the very minimum, if data can be selected to cohere with a theoretical model (e.g., the latent variable model for neuroticism), then at least the theory resists falsification by being empirically adequate.

Throughout the history of science, the validity of new scientific instruments has been subject to debate, often because there is no independent way to assess what the instruments are indicating. When Galileo claimed that his telescope revealed four moons around Jupiter, it was not possible travel to Jupiter to see if he was correct. Similar concerns arose with the introduction of microscopes and particle accelerators—and more recently with factor analysis and its emphasis on latent variables, realistically interpreted.

Borsboom (2008) specifically questions the assumption that the results of factor analysis indicate the existence of latent variables as endogenous entities that cause the patterns of thought, emotions, and behaviors that are encoded in item responses. In an earlier article Borsboom and his colleagues argued that realist assumptions (as opposed to instrumentalist and constructivist assumptions) are inherent to the logic of latent variable models, but constructing latent variables as real causal variables in persons is not among those assumptions.
PERSONALITY DISORDER AND VALIDITY: A HISTORY OF CONTROVERSY

(Borsboom et al. 2003). According to Borsboom and colleagues, we cannot justifiably say that Woody Allen’s obtaining of a high score on our neuroticism scale was the causal result of his possessing a high dose of the latent variable N. What we can justifiably say is that people whose position on N is similar to Woody Allen’s tend to obtain high scores on the neuroticism scale. It is like saying that people who earn a high score on a college admissions tests tend to succeed in college but not like saying that a latent variable called college admissions aptitude caused Bill Clinton to do well in college.

In this view, once all the internal factors that lead Woody Allen to obtain a high score on the neuroticism scale are listed, we would not gain additional information by adding “Neuroticism” to the list. The latent variable of N is not an endogenous causal entity in the scientific realists sense, but an emergent population property that accounts for individual differences. This is similar to a point made in Meehl’s (1954) exploration of clinical and statistical prediction: Statistics provide information about population probabilities.

How then are we to understand individuals? Borsboom argues that an alternative model would explain patterns (either personality or personality disorders) by means of direct causal relationships between emotions, thoughts, and behaviors. In this view, the symptoms (fallible indicators) are not the result of a hidden causal essence behind the pattern (the latent variables such as neuroticism and antagonism), but are directly causally related to each other. For example, the borderline personality symptom “affective lability” may causally influence the symptom of “identity instability,” and an unstable identity may recursively help maintain a labile affective temperament. According to this kind of “network model,” borderline personality is a name for a casual system rather than a manifestation of a latent structure in the head.

As Borsboom et al. (2003) note, however, these competing models are subject to empirical test. For these same reasons, Krueger et al. (2010) call causal network model theorists to task for not directly comparing network models to latent variable models, and suggest one reason that they do not do so is because the type of model they propose is not well suited to achieve goodness of fit because it admits too many variables. Essentially, network models are relatively lacking in parsimony because they posit numerous direct connections among specific and narrow symptoms, as opposed to positing more unifying theoretical principles to help account for numerous specific empirical observations. Some important virtues of more parsimonious latent variable models, Krueger et al. claim, is their heuristic value for generating progressive research program (in a Lakatosian sense) and their potential for organizing the domain in what Phillip Kitcher (2001) calls a “well-ordered” way.

To conclude this section, there is some reason to doubt every one of the claims that five factor model advocates make to justify their scientific realism about personality traits. The five factors were not simply discovered, they may not represent the fixed, universal structure of personality, they are biologically based but in a complex dynamic way, and it is not clear that latent variables must be conceptualized as endogenous causal entities.

Conclusions

The concept of personality disorder, from its very inception, has been infused with philosophical assumptions and controversies. We showed how personality became a secularized
version of character in the medical community, and how the concept of personality disorder developed beyond is roots in degeneration theory and the disease model. The problems of the relationship between personality and issues of character and of what makes personality "pathological" still persist. The issue that is likely of most concern to conceptually minded scientific psychologists, that about the ontological and causal status of latent traits, are also the most philosophically difficult issues reviewed in this chapter. With respect to the problem of latent traits, the conceptual problems that are of concern to psychologists and psychiatrists are difficult for outsiders to grasp, as are the concerns of philosophers. We realize that this means that there may be no natural audience that is interested in solutions that integrate these various sets of concerns, but contend anyway that a variety of research programs in psychology and psychiatry could benefit from philosophical attention to these problems.

REFERENCES


Watson, R., Deary, I., and Austin, E. (2007). Are personality traits items reliably more or less 'difficult'? Mokken scaling of the NEO-FFI. *Personality and Individual Differences, 43*, 1460–9.


