Philosophy of psychiatry: Theoretical advances and clinical implications

Dan J Stein[1], Kris Nielsen[2], Kenneth S Kendler?[3], Anna Hartford[1], Anne-Marie Gagné-Julien[4], Shane Glackin[5], Karl Friston?[6], Mario Maj[7], Peter Zacher[8], Awais Aftab[9]

[1]SA MRC Unit on Risk & Resilience in Mental Disorders, Dept of Psychiatry & Neuroscience Institute, University of Cape Town; [2]Dept of Psychology, Victoria University of Wellington, Wellington, New Zealand; [3]Departments of Psychiatry, and Human and Molecular Genetics, Virginia Institute of Psychiatric and Behavioral Genetics, Medical College of Virginia/Virginia Commonwealth University, Richmond, VA, USA; [4]Department of Philosophy, Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal (UQAM), Montreal, Canada; [5]Department of Sociology, Philosophy, and Anthropology, University of Exeter, Exeter, UK; [6]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London, UK; [7]Department of Psychiatry, University of Campania "L. Vanvitelli", Naples, Italy; [8]Department of Psychology, Auburn University Montgomery, Montgomery, AL, USA; [9]

Work at the intersection of philosophy and psychiatry has an extensive and influential history, and has received increased attention recently, with the emergence of professional associations and a growing literature. In this forum, we review key advances in work on philosophy and psychiatry, and their related clinical implications. First, in understanding and categorizing mental disorder, both naturalist and normativist considerations are now viewed as important – psychiatric constructs necessitate a consideration of both facts and values. At a conceptual level this encourages moving away from strict scientism to soft naturalism, while in clinical practice this facilitates both evidence-based and values-based mental health care. Second, in considering the nature of psychiatric science, there is now increasing emphasis on a pluralist approach, including ontological, explanatory, and value pluralism. Conceptually, this approach again shifts away from a reductionistic scientism towards a liberal naturalism. Clinically, this view emphasizes the importance of a broad range of "difference-makers", as well as a consideration of "lived experience" in both research and practice. Third, in considering a range of questions about the brain-mind – and how both somatic and psychic factors are key in mental disorders – conceptual and empirical work on embodied cognition provides an increasingly valuable approach. Viewing the brain-mind as embodied, embedded, and enactive provides further support for a conceptual approach to the mind-body problem that facilitates the clinical integration of advances in both cognitive-affective neuroscience and phenomenological psychopathology.

Key Words: Philosophy of psychiatry, naturalism, normativism, scientism, reductionism, values-based care, pluralism, mind-body problem, embodied cognition, enactivism

Work at the intersection of philosophy and medicine makes an important contribution by considering key metaphysical issues (e.g., what is the nature of disease?), epistemological questions (e.g., what is evidence-based practice?), and ethical matters (e.g., what is good clinical practice?). Analogous questions arise at the intersection of philosophy, psychiatry, and clinical psychology. Since ancient times, implicit and explicit responses have had a crucial influence on clinical practice. In the West, for example, Aristotle developed a concept of the "golden mean" and related ideas about psychopathology, while in the East Chinese philosophers developed concepts about yin and yang – ideas that have since influenced clinicians, and that continue to do so (Pearson, 2018; McLeod, 2021).

Advances in science after the Enlightenment raised anew conceptual questions about medicine and psychiatry. Karl Jaspers is a particularly seminal figure in the history of philosophy of psychiatry; he not only wrote a key textbook of clinical psychiatry ("General Psychopathology"), but he advanced ideas about best to conceptualize and research mental disorders (Schwartz, Mokalewicz and Wiggins, 2017). His approach has had an enduring and substantial influence on clinical concepts and practice (McHugh and Slavney, 1998). In recent decades, these questions have received increasing attention, with the emergence of professional societies and conferences, as well as journals and textbooks specifically devoted to philosophy and psychiatry (Fulford, Thornton and Graham, 2006; Sadler, van Staden and Fulford, 2015; Aftab and Waterman, 2021; Tekin and Bluhm, 2022).

An influential literature has emphasized the various competencies that health care professionals should acquire (Sidhu *et al.*, 2023). More recently the notion of "conceptual competence" has been proposed; in health care conceptual competence refers to "the transformative awareness of the ways by which background conceptual assumptions held by clinicians, patients, and society influence and shape aspects of clinical care" (Aftab and

Waterman, 2021). These assumptions relate to a range of issues including concepts of disease, professional values, causal explanations, and the mind-body problem. Here, we aim to foreground the importance of conceptual competence for psychiatry.

In the health care sciences, there has been growing attention to evidence-based approaches, and state-of-the-art-reviews are expected to synthesize the literature in a rigorous way (Barry, Merkebu and Varpio, 2022). In philosophy, however, there is ongoing debate not only about the parameters of good philosophy, but also about whether the field actually makes progress over time (Blackford and Broderick, 2017; Williamson, 2021). In this paper, we focus on three areas at the intersection of psychiatry and philosophy. These exemplify a broad range of conceptual debates in the field: they are areas in which there appears to have been progress – if not in resolving all conceptual issues then at least in articulating them clearly – and have particular relevance for clinicians.

We begin by considering responses to the key question of the nature and boundaries of psychopathology, an issue that has long been at the core of philosophy of psychiatry. We then move on to consider questions about the nature of psychiatric constructs and explanations in general, and about how best to think about the brain-mind in particular. In outlining the advances that have been made, and their clinical implications, we argue that there has been a growing and useful emphasis in the field on soft naturalism, on explanatory pluralism, and on embodied cognition, concepts that we will explore in more detail.

THE NATURE OF "DISORDER" AND THE INTERPLAY OF FACTS AND VALUES

In the latter part of the 20th century, a group of thinkers, often referred to as neo-Kraepelinians, saw themselves as ending the dominance of psychoanalysis, countering the

antipsychiatry critique of the field, and re-orienting psychiatry with the mainstream medical tradition (Klerman, 1978). In their view, late 19th century European psychiatry became aligned with the rest of medicine when Emil Kraepelin proposed an influential classification of mental diseases based on clinical description and natural history of the illness. Likewise, the neo-Kraepelinians claimed that precisely defined and reliable diagnostic criteria could be used to discover the specific biological causes of psychiatric syndromes and establish psychiatry as a branch of medicine (Andreasen, 1985; Guze, 1992).

Although Robert Spitzer, the architect of the DSM-III, had some key differences with the neo-Kraepelinians, their ideas helped to undergird the development of DSM-III (Wakefield, 2022). Furthermore, advances in psychopharmacology in the 1960s helped support a view that psychiatric disorders were discrete entities with specific pathophysiologies, and so would respond differentially to medication. Indeed, Donald Klein, a psychopharmacologist whose work influenced the development of DSM-III, put forward the notion of "pharmacological dissection". He held that not only did mental disorders respond selectively to particular medications, but that pharmacological dissection might be useful in delineating different forms of a particular condition; thus atypical depression responded preferentially to monoamine oxidase inhibitors (Klein, 1964; Quitkin, Rifkin and Klein, 1979).

In the 21st century, however, the relationship between DSM and biological psychiatry has shifted, with biologically-oriented psychiatrists emerging as prominent critics of the DSM. Thomas Insel, when director of the NIMH, exemplified this shift. He emphasized that psychiatric disorders are brain circuit disorders, and that descriptive diagnoses – based on symptoms rather than laboratory tests – are not in alignment with the rest of medicine (Insel and Quirion, 2005). Further, given that DSM categories are not biologically-based, the use of DSM categories in research interferes with rather than promotes the discovery of biological

causal mechanisms of psychopathology. Hence Insel proposed, in place of the Research Diagnostic Criteria that were key to the development of DSM (Kendler, Muñoz and Murphy, 2010), a set of Research Domain Criteria, which emphasize translational neuroscience as providing a foundation for psychiatry (Cuthbert and Insel, 2013).

Indeed, it might seem that over the course of its history, psychiatry has lurched from one model to another, in which entirely different concepts of mental disorder prevail. In the United States, it is notable that psychoanalytic thought held sway for many decades, before giving way to a more neurobiological perspective (Luhrmann, 2001). Even within more biologically-based psychiatry, there has been considerable debate about the fundamental nature of mental disorders, as perhaps most notably exemplified by the decision to omit homosexuality from the DSM (Bayer, 1987; Zachar and Kendler, 2012). In philosophy of science Kuhn's notion of scientific paradigms has been enormously influential (Kuhn, 1962), and in psychiatry many have proposed paradigm shifts for the field (Stein *et al.*, 2022).

The clash between different psychiatric models has brought conceptual issues to the fore. The notion put forward by both neo-Kraepelinians and translational neuroscientists that mental disorders are brain disorders, for example, raises a series of inter-related and perennial philosophical debates including the nature of mental illness, the relationship between different explanations of mental illness, and the mind-body problem (Maj, 2013). Each of these conceptual debates has important clinical and research implications, as exemplified in vociferous debates between biologically-oriented psychiatry and psychoanalysis, between those who emphasize intrinsic causes vs social determinants of mental illness, or in calls by NIH to move from DSM to RDoC. These conceptual debates seem increasingly urgent given the growing recognition of the burden of mental disorder, and the ongoing need for better interventions.

In this section we will focus on the nature of mental illness; this provides a foundation from which to consider other key conceptual debates as the paper proceeds. The nature of mental disorders in turn raises a series of subsidiary questions, each of which will be addressed here in turn:

i) What justifies the position that a particular biological or behavioral state is a disorder? Are our judgments based on objective natural facts, or are they merely normative and reflective of our interest?

ii) How should deviations from health be classified? Do diagnostic distinctions reflect objective, natural features of psychopathology or do they reflect our clinical and scientific interests?

iii) Are mental disorders best considered as universal entities that are similar across individuals, or as shaped in particular ways that are unique for each person? What are the implications of psychiatric diagnosis for personal agency?

Disorder Status: Naturalism and Normativism

In a straightforward binary version of this debate, naturalism and normativism are opposite and diametrically opposed views (Table 1). The phrase "the disorder wars", comes to mind (Bortolotti, 2020). On one end lies *strong naturalism* which is the view that the concept of "disorder" can be described in completely factual and value-free terms and can be best studied using the methodologies continuous with those used in the natural sciences such as chemistry and genetics. Many biological psychiatrists of the late 20th century held this view to the extent that they accepted psychiatric disorders to be caused by neurobiological

dysfunction and understood abnormal psychology to be the result of objectively deviant brain functioning (Andreasen, 1985; Guze, 1992).

In the philosophy of medicine, the notion of disorder as objective deviation from a state of health is notably expressed by Christopher Boorse. For Boorse health is a state of normal biological functioning and functions are normal if they make a causal contribution to survival or reproduction that is typical for the species (Boorse, 1977). Boorse has been remarkably persistent in maintaining this view; twenty years after his original papers he published a lengthy rebuttal to his critics (Boorse, 1997), and nearly two decades later at a symposium on his work he again countered his critics (Boorse, 2014). Indeed, it has been suggested that after Boorse, philosophers of medicine must either work within his theory or explain why not (Schramme, 2014).

On the other end of the divide *strong normativism* holds that there is no natural, objectively describable set of biological processes that we can characterize as "dysfunctional" and hence disorder attributions are thoroughly value-laden. Normativists differ on the nature of these value judgments. For Bill Fulford, a seminal author in modern philosophy of psychiatry, disorder is inherently normative because it is grounded in the "illness experience," the patient's direct experience of something having gone wrong which is dependent upon social or folk-psychological intuitions of what is abnormal (Fulford, 1990). For Fulford the value-ladenness of the illness experience not only unites medicine and psychiatry, but also humanizes both fields.

Thomas Szasz, renowned for his critique of psychiatry, provides an entirely different view from Fulford. For Szasz, disorder judgments in psychiatry are judgments of deviance based on sociocultural norms, with no evidence of the presence of a biological disease. Szasz's view of psychopathology arises from a strong naturalist view of physical disease

together with a strong normativist view of mental illness: physical disease is value-free, and if states we call "mental disorders" are value-laden, then their characterization as disorders or illnesses is a category error, a myth. For Szasz, value-ladenness becomes a reason to question the medical legitimacy of psychiatry because illnesses ought to be described in terms of objective, pathological changes.

It is crucial to appreciate, however, that those who view disorder concepts as inherently value-laden do not necessarily deny the biological reality of the afflictions. Naturalists and normativists can both agree on the physiological and behavioral facts at hand and yet may disagree on whether the state in question is healthy or disordered (Glackin, 2019). As the philosopher Rachel Cooper has illustrated using a weeds and daisy metaphor, we can all agree on what a daisy is as a species, but disagree on its status as weed (Cooper, 2004). Similarly, researchers can agree on the biological mechanisms of premenstrual dysphoric disorder, but disagree on its status as a psychopathology (Gagné-Julien, 2021b; Stein, Palk and Kendler, 2021).

For naturalists, medicine is at its theoretical core a scientific discipline like other natural sciences and subject to a similar sort of interplay of natural facts and human interests (Boorse, 1997). For normativists, disorder concepts are not fundamentally scientific but rather are clinical and practical concepts. They are grounded in the experiences of distress, disability, and disruption that are interpreted to indicate that something has gone wrong and that lead patients to seek professional help for their problems. From a normativist perspective, medicine is at its core a practical activity aimed at reducing human suffering and enhancing well-being (Pellegrino and Thomasma, 1981; Fulford, 1990; Glackin, 2016).

The naturalist-normativist debate acquires a particular valence in psychiatry in part because of the way value-ladenness has been wielded by antipsychiatry figures, such as Szasz,

to challenge the notion of mental illness. New critical movements have gone beyond the Szaszian approach, by exploring how social and cultural values impact views of the normal and the pathological. Neurodiversity studies, for example, argue that cognitive profiles such as autism may be socially disabling, but are not intrinsically pathological (Rosqvist, Chown and Stenning, 2020; Chapman, 2021). Some studies similarly resist the pathologizing of diversity and emphasize social factors as a cause of distress (LeFrançois, Menzies and Reaume, 2013; Beresford and Russo, 2022).

Binary positions have the advantage of being straightforward. One disadvantage is that when they are understood in opposition to each other, their differences are often accentuated, such that each position may be defined by what the other rejects. Further, an important development in philosophy of science has been an appreciation of the role values play in science and recognition that the notion of value-free science is not only untenable but also undesirable (Longino, 1990). For example, values influence which scientific problems are prioritized, how they are studied, how uncertainty is managed, how much evidence is considered sufficient, and how scientific evidence is used to inform practical decision-making. The incorporation of values and human interests into a broader notion of scientific objectivity has enriched our understanding of natural sciences.

Strong naturalism runs the risk of scientism – of over-reliance on what is currently perceived as factual (Boudry and Pigliucci's, 2017; de Ridder, Peels and van Woudenberg, 2018), while strong normativism runs the risk of a relativism where knowledge is ignored and where "anything goes" (Feyerabend, 1975). In philosophy, a position that has been termed "soft naturalism", attempts to avoid both scientism and relativism, and to acknowledge the importance of both facts and values in science (De Caro and Macarthur, 2022). And in philosophy of psychiatry a number of different proposals have been put forward as to how

best to incorporate both naturalist and normative considerations in conceptualizing mental disorders (Amoretti and Lalumera, 2022).

A particularly influential integrative position, Jerome Wakefield's harmful dysfunction analysis, is a hybrid view that combines both naturalism and normativism in roughly equal measures (Wakefield, 1992; Spitzer, 1999). One component of disorder, "dysfunction," is defined in value-free, evolutionary terms. Dysfunction refers to the failure of biological or psychological mechanisms to perform the function which they were naturally selected to perform. The second component of disorder, harmfulness, is normative, and in Wakefield's view, largely determined by social standards. Wakefield has applied his harmful dysfunction analysis to a broad range of psychiatric disorders, and like Boorse has engaged widely with critics over several decades (Faucher and Forest, 2021).

According to Wakefield, for instance, developing depression in reaction to a stressor such as loss is an evolutionary designed adaptive response to adversity and not a dysfunction; the DSM, therefore, makes an error by classifying such depressive reactions as disorders. It is only when depression occurs out of the blue, or does not resolve once the stressor is no longer active, or accompanied by complicated features (such as suicidal ideation, psychosis, and psychomotor retardation), that it becomes reasonable for us to assume that mechanisms designed to regulate sadness in response to loss and adversity have failed (Horwitz and Wakefield, 2007; Wakefield and Schmitz, 2014).

One recent alternative to Wakefield's account is a hybrid account offered by Jon Tsou. Tsou defines mental disorders as biological kinds (value-free component) with harmful effects (normative component), and by doing so, bypasses speculation about what normal psychological functions are products of natural selection (Tsou, 2021). Instead, drawing on the work of Richard Boyd on how natural kinds reflects a cluster of underlying properties

(Boyd, 1999), he argues that valid biological kinds are those that exhibit characteristic regularities due to stable sets of interacting biological mechanisms, which allows us to make inferences and predictions about diagnostic categories. We can do this because the properties that define scientifically valid kinds are produced by similar sets of causal mechanisms.

For Tsou, schizophrenia is a disorder because it entails shared causal mechanisms that result in an identifiable cluster of properties with predictable regularities (i.e., it is a biological kind) and because it compromises the capacity of a person to function adequately as judged by sociocultural standards (i.e. they are harmful). However, Tsou would also include as disorders normal psychological reactions to stress, such as acute depression, which are characterized by biological mechanisms that fall in the normal range of function. Thus, the naturalistic standard of being a biological kind is broad enough to accommodate the range of conditions that mental health professionals treat.

Additional ways of bridging the naturalist-normativism divided have been proposed, including by the authors of this review (Stein *et al.*, 2010; Stein, 2013; Zachar and Kendler, 2017; Nielsen and Ward, 2020; Conley and Glackin, 2021; Gagné-Julien, 2021a). Gagné-Julien, for example, argues that judgments about dysfunction are not merely objective, but are in fact value-laden. She suggests that provided that appropriate procedures are in place for making judgments about disorders, these can be socially objective (Gagné-Julien, 2021a). Nielsen and Ward argue that the key norm violation for disorders is breakdown in the norms that support an individual's functioning within their social context (Nielsen and Ward, 2020). They attempt to "naturalize normativity" by noting that in the psychiatric domain, disorders entail cognitions and behaviors that run counter to an individual's self-maintenance and

adaptation needs; disorder status is therefore based on the needs of the individual, rather than on societal norms.

Strong naturalism can be tempered by acknowledging that values and human interests play important roles in clinical and scientific contexts. Many would agree that the concept of disorder invokes value-laden notions such as disability, harm, and suffering (Fulford, 2001; Sadler, 2004; Fulford *et al.*, 2005). Authors such as Lawrence Reznek, Dominic Murphy, and Rachel Cooper consider disorders to be natural processes that are held together in virtue of human interests, akin to categories such as “weed” or “vermin” (Reznek, 1991; Cooper, 2004; Murphy, 2006). Such weaker forms of naturalist concepts of disorder may be seen as exemplars of a soft naturalism that emphasizes the complexity and fuzziness of the world, as well as the need to address both the mechanisms underlying disease as well as the experience of illness (Stein, 2021).

A view of science as influenced by values can also provide nuance to strong normativism. Thus, a strong version can be qualified by insisting that the values that influence our definition of mental illness can be discussed and critiqued to reach a consensus on the type of values that are desirable in psychiatry (e.g., values concerning human flourishing, well-being, harm reduction, vs. oppressive values such as racism and sexism). Notably, Spitzer was open to articulating the values underpinning DSM-III (Spitzer, 2001). Further, several authors have advocated for consultative decision-making processes that would include patients’ voices on the question "what is a mental illness?” so at to ensure that patients’ interests are represented in psychiatric concepts and classifications (Stein and Phillips, 2013; Fulford, Bortolotti and Broome, 2014; Bueter, 2019; Fuss *et al.*, 2019; Gagné-Julien, 2021b; Friesen and Goldstein, 2022; Knox, 2022; Tekin, 2022; Rose and Rose, 2023).

Strong normativism can also be tempered by acknowledging that broad scientific agreement can be achieved on the co-occurrence and co-variation of signs and symptoms that characterize the psychiatric conditions labelled as disorders. For example, whether or not people have the symptoms of anorexia nervosa can be seen as an empirical matter and the decline in functioning associated with these difficulties can be recognized by all observers regardless of the value-laden nature of the standards by which functioning may be judged to be impaired. Furthermore, scientific agreement can also be reached on the involvement of particular neurobiological processes in particular psychiatric conditions (Hyman, 2021), even though these processes may not be characterized as "dysfunctional" on neuroscientific grounds alone (Conley and Glackin, 2021).

Pragmatic considerations have assumed an increasingly prominent role in the conceptualization of mental disorders. Pragmatic accounts focus on the importance of human interests, but they tend to focus on clinical and scientific goals rather than sociocultural norms and values. For instance, in articulating the notion of a practical kind, Zachar argues that the development of disorder concepts in the DSM and the ICD can be seen as an attempt to calibrate concepts to multiple goals such as enhancing reliability, supporting etiopathological validity, facilitating communication, guiding treatment, minimizing stigmatization, and promoting future research (Zachar, 2015).

The bridging of the naturalist-normativist divides provides key lessons for clinicians. In particular, such bridging provides an important foundation for complementing evidence-based care with value-based care. Evidence based care is largely focused on synthesis of the medical literature, while values-based health care reminds us of the importance of assessing and addressing patient's values. Values-based care is consistent with a model of patient-centered practice, where the values of individual patients are central to evidence-based

clinical decision-making. Fulford's model of values-based health emphasizes that it is skills-based, with the most important skills being awareness (of values), reasoning and knowledge (about values), and communication skills (Fulford, 2008). Each of these skills draws on philosophical sources, but also exemplifies good psychiatric practice.

Psychiatric Classification: Moving Beyond Essentialism

Once we have implicitly or explicitly identified a class of mental disorders, a set of psychopathological states, or a community of psychiatric conditions/mental health problems, we can further ask: How do we distinguish between conditions within the class of mental disorders? How do we map the territory of psychopathology? How do we differentiate one mental health problem from another, and how do we demarcate disorder from normality?

Philosophy of psychiatry has been helpful in clarifying the metaphysical and methodological assumptions that guide the search for answers to these questions. These include the nature of disorder categories, assumptions pertaining to the methodological approaches of description and classification, and clarity on the use of diagnostic operationalization. A common metaphysical assumption in psychiatric classification has been essentialism. Essentialism is the notion that categories have essences, identity-determining properties that all members have in common and that distinguishes them from members of other categories. Essentialism is linked to ideas about natural kinds, a categorization that reflects the structure of the natural world, and in the context of psychopathology, is linked to notions of psychiatric disease entities that are to be discovered through scientific inquiry, "carving nature at her joints", as Plato put it (Campbell, O'Rourke and Slater, 2011; Kincaid and Sullivan, 2014).

Philosophy of biology and of psychology have recently focused on how biological and mental processes and causal mechanisms undergird observed phenomena (Bechtel, 2007; Craver, 2007; Nicolson, 2012). When causal processes and mechanisms are well-understood, medical professionals are often able to use them to categorize conditions. This is particularly the case when discrete disease entities can be identified, such as in infectious diseases where classification is based on identification of the causative pathogen. However, when the processes and mechanisms of an illness are complex, dimensional, or multifactorial, knowledge of etiology by itself doesn't offer an optimal classification and we have to rely on additional considerations – on what we want the classification to accomplish – to draw boundaries and set thresholds. This applies to many areas of medicine, including internal medicine, neurology, and rheumatology, but is an issue that is more pervasive and pronounced in psychiatry (Glackin, 2017; Huda, 2019).

From a somewhat simplified metaphysical perspective, we may think of a classification as demarcating natural kinds, practical kinds, or social kinds. If psychiatric classifications such as DSM and ICD were classifying natural kinds, we'd expect each diagnosis to correspond to an entity that exists in the structure of the world, independent of human interests (Campbell, O'Rourke and Slater, 2011; Kincaid and Sullivan, 2014). This assumption has been important to the Kraepelinian and neo-Kraepelinian tradition of psychiatric classification. Emil Kraepelin, for instance, believed in the existence of natural disease entities in psychiatry, and in addition held the view that pathological anatomy, etiology, and clinical symptomatology including course of illness would all coincide in the case of natural disease entities (Heckers and Kendler, 2020).

The assumption that there are natural disease entities in psychopathology was also adopted by the neo-Kraepelinians, and implicitly guided the development of DSM-III

(Blashfield, 1982; Aftab and Ryznar, 2021). Furthermore, the Kraepelinian notion of convergence of validators was also accepted by Robins and Guze, who assumed that their proposed validators of clinical description, laboratory findings, course of illness, and family studies would all point towards the disease entities (Robins and Guze, 1970). This set the agenda for a research program for the next several decades in which researchers sought to validate the DSM diagnostic constructs.

By the 1990s, however, there was growing recognition that different validators of psychiatric classification might not inevitably align to offer a single privileged classification – perhaps like the periodic table of elements (Kendler, 1990). Rather, different validators suggest alternative mappings of the space of psychopathology (Solomon and Kendler, 2021). For example, in the study of schizophrenia, shared family history suggests a broad mapping (schizophrenia spectrum), whereas poor outcome indicates a narrower mapping (schizophrenia). In such a scenario, empirical facts do not determine which validators we ought to use. Our choice of validators depends on what we value more – which could differ from practitioner to practitioner and from context to context.

In opposition to the natural kind view has been the skeptical view that the categories of psychiatric classifications are social kinds, almost entirely constructed by social processes. This view appeals to many critics of psychiatry, who point towards the obvious influence of sociocultural factors on the presentation of psychiatric conditions and the inability of psychiatric research to identify valid biomarkers. The socially constructed view is further supported by examples such as “hysteria” and “multiple personality disorder” whose popularity among clinicians at various points in history has resembled the rise and fall of fashions. There is also increasing awareness that psychopathological phenomena are subject

to “looping effects,” such that the very act of classification modifies the behavior of the individual classified, further supporting the social constructionist view (Hacking, 1995).

However, the social kinds view in its extreme articulation seems untenable as it fails to take into account that scientific research has discovered relationships between neurobiological processes and psychiatric symptom clusters, albeit these relationships do not necessarily correspond to specific DSM or ICD categories. For instance, psychiatric research has identified hundreds of genetic variations that are associated with a range of psychiatric disorders, so that genetic influences on psychopathology often cut across DSM diagnostic boundaries (Smoller *et al.*, 2019; Andreassen *et al.*, 2023). The relationship between genetic variants and psychopathology is therefore complex and transdiagnostic, but not absent or chaotic (Kendler, 2013).

The notion of practical kinds offers a different position that aligns with the soft naturalist view that psychiatric science is both a scientific and social process. There may be no “natural joints” in psychopathology but there are scientific facts in the form of symptom patterns and co-variation that constrain any scientific attempts at nosology (Zachar, 2014). Within those constraints, the boundaries we draw will often reflect our pragmatic goals, and diagnostic thresholds will be influenced by both facts and values. Practical kinds are useful heuristic constructs that categorize the neurophysiological and psychological space in ways that serve our scientific and clinical goals. The pragmatic nature of psychiatric classification is also supported by considering the history of psychiatric nosology - this shows the contingent nature of our contemporary diagnostic constructs and how our classifications would have looked quite different had certain key historical figures in psychiatry not existed or had made different choices (Kendler and Zachar, 2015; Kendler, 2016a).

Giving up essentialist assumptions about natural kinds in psychopathology allows us to appreciate the complexity of psychopathology and makes it possible for us to map and model psychiatric phenomena using different approaches. Instead of diagnosing symptom clusters, for example, idiographic approaches focus on the uniqueness of the individual psychiatric patient – how their mental health problems arise from a specific combination of predisposing factors, developmental history, life experiences, behavioral adaptations, and psychological defense mechanisms. Such an approach utilizes broad principles of psychobiological functioning to formulate a narrative specific to a patient. The aim of classification, then, is to aid the development of the clinical formulation.

The failure to identify disease categories has spurred psychometric efforts to model psychopathology. Psychometric analysis goes beyond manifest variables, which can be directly measured or observed, to mathematically model latent or hidden variables, which are factors that cannot be directly observed and only emerge through statistical analysis. With a few exceptions, the mathematical models have better fit with latent dimensions rather than latent categories (Haslam *et al.*, 2020). What this means is that rather than psychotic disorder being a category with members who are qualitatively distinct from those who are not psychotic, psychoticism refers to a dimension of functioning on which everyone has a position. The differences between low and high scores are quantitative not qualitative (Kessler, 2002).

This quantitative statistics research program has achieved some success in recent years in the form of the Hierarchical Taxonomy of Psychopathology (HiTOP) consortium (Kotov *et al.*, 2017), which has relied on large datasets to identify a hierarchy of dimensions ranging from the most homogenous (symptoms and traits) to the least homogenous (spectra and super-spectra). The spectra such as “Internalizing” and “Externalizing” represent broad

psychopathological tendencies, and may possess a stability and convergence that is lacking at the syndromic level.

That said, a range of conceptual issues have been raised about the psychometric approach (Haeffel *et al.*, 2022). First, in clinical practice there do seem to be some discrete entities, which respond to specific treatments; narcolepsy, for example, can be diagnosed using an accurate biomarker, and can be effectively managed using particular medications. Second, dimensions and categories are more interchangeable and potentially complementary than is sometimes articulated (Kessler, 2002; Fulford and Handa, 2018). And of particular relevance to positions that emphasize the importance of causal mechanisms for classification, psychometric approaches emphasize descriptive features and may elide underlying etiology (Zachar and Kendler, 2017).

Another strand of philosophical inquiry has focused on the use of operational definitions employed by the DSM. In an effort to improve inter-rater reliability and to facilitate psychiatric research, the DSM from its third edition and onwards has offered operationalized criteria for each disorder that specify details such as a list of (relatively specific) symptoms, number of symptoms that must be present, and the duration for which they must be present. For major depressive disorder, for instance, the DSM requires that 5 out of a list of 9 symptoms be present (with one of them being low mood or anhedonia) and that the duration must be at least 2 weeks. How should the relationship between the criteria and the disorder be conceptualized? Lack of clarity in this regard leads to another form of confusion in which operational criteria are thought to constitute the disorder itself.

Operational definitions are partial definitions that do not specify all the details of the phenomena being studied. Operational definitions have an element of vagueness that becomes evident when new scientific questions force us to articulate concepts with greater

precision. DSM criteria excluded nonspecific symptoms from operational criteria (such as anxiety in depression) to improve specificity but these nonspecific criteria are still a part of the depressive syndrome. The polythetic nature of DSM criteria allows for many different symptom configurations to meet disorder threshold, but these different symptom configurations do not constitute different disorders. Instead, these configurations are better understood as different ways in which we can identify a disorder.

Kendler has elaborated on the distinction between diagnostic criteria as indexical and constitutive (Kendler, 2017). When diagnostic criteria are understood to be indexical, they are understood to be fallible ways to identify a disorder; when they are understood to be constitutive, the symptom criteria *are* the disorder. DSM criteria are intended to be indexical, and viewing them as constitutive is a conceptual error. Thus, for example, DSM criteria for major depressive disorder allow for 227 ways for criteria to be met, but these are ways of indexing major depression, not 227 types of major depression (Fried and Nesse, 2015). There is no single and privileged correct operationalization, rather different operational definitions can be refined and optimized for different purposes.

Taken together, an emerging contemporary view of psychiatric taxonomy incorporates the dimensionality of psychopathology (there are no discrete entities), insights from complex dynamic systems (relatively stable symptom patterns can emerge from irreducible interactions between multiple factors), and perspectives from embodied cognition (causal mechanisms traverse the brain, body, and environment). Such a view of psychopathology does not render categorical diagnostic systems such as DSM and ICD invalid or useless, but it forces us to give up the essentialist bias that has led us to reify them - to attribute a correspondence to objective reality that they do not possess (Ghaemi, 2009; Hyman, 2010).

How is this view of taxonomy relevant to clinical practice? In our view clinicians need to strike a balance between being aware of the work that has gone into, and the value of our of our nosology, while also being mindful of its tentative nature and significant limitations. In particular, although DSM has clear clinical utility, it has often been criticized for facilitating a checkbox approach to psychiatric assessment and evaluation. Clinicians ought to be aware that important features of mental disorders may well have been described in the psychiatric literature, and yet may not be listed in DSM (Kendler, 2016b). Further, while diagnosis may begin with DSM or ICD, a comprehensive evaluation needs to assess a range of domains, including clinical subtypes, symptom severity and staging, cognitive schemas, environmental stressors, and protective factors (Maj *et al.*, 2020). Clinical formulations need to supplement our growing knowledge of psychiatric disorders as psychobiological kinds, with an idiographic understanding of each individual patient (Schwartz and Wiggins, 1985).

Psychiatric Diagnosis and Personal Agency

In this section, we will focus primarily on the issue of personal agency, noting that debates concerning psychiatric taxonomy may have important implications for individual self-conception and self-understanding (Tekin, 2014). As noted, DSM criteria are not constitutive of disorders, but they are nevertheless often taken as such, and the influence of the DSM on how mental disorders are perceived is profound. Concern has been raised about the undue extent of this influence, especially given the inevitable neglect of person and context specific factors in generalized diagnostic criteria (Tekin, 2014).

More broadly, debates about the nature and classification of disorders are also implicated in the effort of patients to understand the boundaries of their selves in relation to

their disorders. Given that both psychiatric conditions and psychiatric medications can affect perceptions, desires and feelings—deep aspects of self-experience—ambiguity and uncertainty can arise with regard to where the "self" begins and ends, and how the self is impacted (or compromised) by both illness and treatment (Sadler, 2007; Dings and Glas, 2020). Ambiguity at the conceptual level—concerning the nature of psychiatric disorders—can further compound the experience of ambiguity at the phenomenological level for patients who are "confronted with the vagueness and uncertainty associated with the issue of 'what *is* a psychiatric disorder'" (Dings & Glas, 2020).

Questions concerning the interplay of agency and mental disorders have also been central to debates concerning the relevance (or irrelevance) of mental disorders to assessments of moral responsibility. While psychopathology has often been treated as paradigmatically exempting or mitigating in the literature on moral responsibility, there has been a growing shift to more nuanced assessments—with no generalized inference from the fact of psychopathology to the appropriateness of exemption or mitigation— and a growing emphasis on the need for case-by-case assessments (King and May, 2022). These more nuanced assessments reflect the larger recognition of person-specific and situation-specific factors that affect the manifestation of psychopathology for any particular individual.

A crucial element of this analysis is how the capacities deemed relevant to moral responsibility are impacted in any particular case. Further nuance is necessitated by the recognition of degrees and spectrums with regard to how the relevant capacities are affected. Cases of clear-cut incapacitation present a less complicated theoretical terrain than the many cases where the relevant agential capacities are diminished or deeply compromised, but nevertheless present. In addiction, for example, it is often implausible to speak of blanket incapacitation, given that aspects of choice and deliberation are often involved. A useful body

of work has explored the question of responsibility in this context (Pickard, 2017; Brandenburg, 2018).

More broadly, the question arises of how different ways of conceptualizing psychiatric disorders affect our attitudes towards affected individuals? While it was presumed that more biological conceptions of disorders would reduce stigmatizing attitudes in general, empirical research points to far more complex interactions (Haslam and Kvaale, 2015). These findings align with theoretical concerns regarding the interpersonal and social costs of perceived diminished agency which, while sometimes decreasing perceived responsibility, might simultaneously increase other forms of aversion. Further research has suggested that biological conceptions may ultimately be more stigmatizing for affected individuals (Loughman and Haslam, 2018). A related body of empirical literature concerns the impacts on these conceptualizations on assessments of the appropriateness of punishment, including legal punishment (Tabb, Lebowitz and Appelbaum, 2019).

Moreover, the impact of psychiatric diagnoses on the lives of those diagnosed has encouraged the use of first-person accounts in the elaboration of psychiatric classification. Rooted in the theoretical frameworks of feminist philosophy of science and standpoint theory, several authors have suggested that given their first-person perspective, people with lived experience can usefully contribute to the development of psychiatric classifications; they may be better situated to assess the impact of changing diagnostic criteria on access to care or the potential risk of stigma associated with certain nomenclature issues, or be better able to identify mismatch between diagnostic criteria and lived experiences classifications (Stein and Phillips, 2013; Fulford, Bortolotti and Broome, 2014; Bueter, 2019; Fuss *et al.*, 2019; Gagné-Julien, 2021b; Friesen and Goldstein, 2022; Knox, 2022; Tekin, 2022; Rose and Rose, 2023).

The clinical implications of work on personal agency and moral responsibility have been usefully debated. In clinical relationships a crucial concern is which interpersonal stances concerning the nature of interpersonal moral expectations and demands will be most beneficial to patients, or lead to the best therapeutic outcomes. As opposed to the "participant stance," of robust and reciprocal interpersonal moral expectation (Strawson, 1962), Pickard has proposed the "clinical stance" of "responsibility without blame" (Pickard, 2017). On this view, responsibility can be asserted in crucial respects, even as we excise the affective aspects of blame such as the emotional reactions of resentment and indignation that usually accompany the experience of having been mistreated or wrongfully harmed. Thus, the clinician acknowledges both the extent of agency (in asserting agential responsibility) but also its limits (in denying or tempering moral responsibility).

In contrast, Brandenburg has argued that the affective, emotive aspects of blame can have an important place between clinicians and patients, with "nurturing reproach" an important aspect of the clinician's "nurturing stance" (Brandenburg, 2018). The idea here is that while full-scale resentment would be inappropriate towards people with compromised capacity, reproach may sometimes be appropriate and even therapeutically beneficial, especially insofar as it acknowledges the capacity to "work towards developing or repairing one's own moral abilities". In interviews with clinicians Brandenburg and Strijbos explored some of the dynamics of this stance, and reported that "a number of clinicians mentioned how, by showing that someone affects them, rather than concealing this—they 'take a person seriously'…. See 'the person, in some sense, as an equal'… or simply 'express [that] you respect the person" (Brandenburg and Strijbos, 2020).

PLURALISM IN PSYCHIATRY

We noted earlier that over the course of its history, different schools of thought have characterized psychiatry. In philosophy of science, Kuhn's notions of dominant scientific paradigms that are incommensurable, and of revolutionary shifts in the paradigms that govern normal science (Kuhn, 1962) have become widely employed, and arguably psychiatry provides a useful exemplar of how different paradigms dominate over the course of time. Indeed, critics of psychiatry have argued that the replacement of one psychiatric paradigm by another entails neither scientific progress nor clinical advancement (Harrington, 2019; Whooley, 2019).

At the same time Kuhn has been criticized for his relativism (Richards and Daston, 2016); after all, scientific models can be reasonably compared, and there may be justifiable grounds for replacing one model with another. In psychiatry, although there certainly have been important shifts in theoretical frameworks, it might also be argued that current clinical research and practice incorporate valid aspects of both psychodynamic and neurobiological approaches, as well as concepts and data from a range of other models of psychopathology. Different models may be able to engage usefully, as evidenced by the emergence of neuropsychoanalysis, or by work on how psychotherapy alters neuroimaging. And psychiatry has arguably advanced precisely by incremental integration of a range of valid models (Stein *et al.*, 2022).

Notably, psychopathology seems to involve multiple causes, and it is possible that different psychiatric models shed light on different causes. Philosophers have long emphasized the importance of multi-causality, with Aristotle famously delineated four types of causes. More recently, in a landmark paper, the ethologist Niko Tinbergen's outlined four

causes of animal behavior (Tinbergen, 1963). Two involve more proximate causes - ontogeny (how does the trait develop in individuals) and mechanism (what structures underlie the trait), while two involve more distal causes – phylogeny (how did the trait evolve) and adaptive significance (how does the trait influence fitness). Notably, Tinbergen's questions are a useful framework for delineating different perspectives on disease pathogenesis (Nesse, 2019).

Aristotle was an enormously prescient philosopher-scientist (Leroi, 2014). He outlined different causes of mental disorder (Pearson, 2018), and there are overlaps between his and Tinbergen's approaches to causality (Hladký and Havlíček, 2013). More recently, Jaspers, drawing on German philosophy of science (Feest, 2010), distinguished between knowledge of causal explanations and understanding of meaningful connections (Schwartz, Mokalewicz and Wiggins, 2017). Indeed, Jaspers can be understood as a methodological pluralist, with his pluralism influencing a range of subsequent authors (Ghaemi, 2007). In their influential text, for example, McHugh and Slavney identified four explanatory methods in psychiatry: diseases, dimensions of personality, goal-directed behaviors, and life stories (McHugh and Slavney, 1998).

In contemporary philosophy of science there is ongoing debate about whether and how diverse explanations can be integrated (Ludwig and Ruphy, 2021). In the 1970s, George Engel, an American internist and psychiatrist, argued that the dominant model of disease was biomedical, and this neglects the psychological and social dimensions of illness (Engel, 1977). He therefore proposed a biopsychosocial model, aiming for a framework that could be used in research, teaching, and clinical care. Clearly, it is important that we avoid both a brainless psychiatry and a mindless psychiatry (Lipowsky, 1989), steering clear of both scientism and

culturalism (which are overly reductionist about science and culture respectively) (Chemla and Keller, 2017; de Ridder, Peels and van Woudenberg, 2018).

However, the biopsychosocial model has also received stinging criticism for being overly eclectic and non-specific (Ghaemi, 2010), and for offering no particular framework to conceptualize multi-level causal interactions (Bolton and Gillett, 2019) or optimal selection of causal mechanisms (Maung, 2021). Further, its practical use in psychiatric formulation has led to an inadvertent reification of "biological", "psychological," and "social" as distinct ontological domains (Aftab and Nielsen, 2021). Ongoing efforts to understand the nature of causal explanations in science in general (Woodward, 2003; Pearl and Mackenzie, 2018), and in psychiatry in particular (Stein, 2008; Kendler, 2012) therefore remain crucial.

An explicit emphasis on pluralism is a relatively recent development in philosophy of science (Kellert *et al.*, 2006). Nevertheless, as evidenced by Aristotle work on multi-causality, it has long historical roots. Notably, Moritz Schlick, the founder of the Vienna Circle, wrote that "every sensible and philosophically honest worldview must be pluralistic. For the universe is variegated and manifold, a fabric woven of many qualities no two of which are exactly alike" (Schlick, 1985). Similarly, American pragmatists such as William James and John Dewey early on emphasized pluralism (James, 1907; Dewey, 1938). More recently, the Stanford School, including philosophers such as Nancy Cartwright and John Dupré, has promoted a pluralist program (Dupré, 1993; Cartwright, 1999).

Work on pluralism has not only been important within philosophy, but also provides a tremendously useful resource for clinicians. Unsurprisingly, for philosophers who regard pluralism as important, there is not a single unified approach to pluralism. Instead a number of different pluralisms have been delineated and developed, and the extent to which an integrative pluralism is possible remains under debate (Ludwig and Ruphy, 2021). We next

consider three important notions of pluralism - *ontological pluralism, explanatory pluralism, and value pluralism* - as well as some of their clinical implications, before going on to consider a number of tools that may be useful in thinking about psychiatric explanations.

Different Notions of Pluralism

First, consider *ontological pluralism*. As noted earlier, the notion of "natural kinds" reflects the possibility that nature can be carved up in an objective way to form discrete entities (Campbell, O'Rourke and Slater, 2011; Kendig, 2016). Exemplars of such natural kinds are often found in physics or chemistry; the periodic table of elements is a particularly compelling one. Ontological pluralists have, however, argued that there are different ways of dividing reality, reflecting different scientific interests and values, and that a range of different classifications may be valid. From the time of Aristotle, pluralists have often looked to biology, where while species can potentially be divided on the basis of their evolutionary history, there are alternative equally valid ways of classifying organisms (Henry, 2011; Dupré, 2012).

Our earlier discussion of mental disorders emphasized that mental disorders are not simply natural kinds that emerge from empirical investigation. At the same time, our constructs of mental disorders are not merely conventional. Instead, our constructs of mental disorder are rigorously informed by scientific research, including work on a range of different validators, which reflect the involvement of a range of different underlying structures and mechanisms. They may be regarded as "soft natural kinds"; although mental disorders cannot simply be discovered by carving nature at her joints, and although our classifications and descriptions of mental disorders are value laden, these entities nevertheless incorporate an

accumulating scientific appreciation of psychobiological structures, processes, and mechanisms (Stein, 2022).

The notion of "soft natural kinds" may be useful in the clinic in a number of ways. Consider the construct of "behavioral addiction". From a neo-Kraepelinian perspective, the classification of substance-use disorders together with gaming and gambling suggests that conditions have overlapping phenotypic features, and share key validators, such as clinical course. In fact, however, the situation may be much "fuzzier": after all the psychobiology of alcohol dependence is likely to differ significantly from that of gambling, given that alcohol has direct toxic effects on the brain. Indeed, a key rationale for lumping these conditions may instead be that from a public health perspective, these conditions all respond to harm reduction interventions (Stein *et al.*, 2018).

Consider also the boundaries between disorder and normality (Geert, Lara and Rico, 2017). Current versions of DSM and ICD appropriately emphasize that the boundary between disorder and normality is not hard and fast, but rather can be a fuzzy and indeterminate one. In some other areas of medicine, biomarkers can be helpful in making the clinical decision as to whether a disorder should be diagnosed, but this is not often the case in psychiatry. Critics of psychiatry may conclude that mental disorders are entirely a matter of convention, and that psychiatric diagnosis is merely a matter of "labelling". However, this ignores the complex reality of mental signs and symptoms: psychiatric phenotypes are not elements in a period table but rather are comprised of overlapping dimensions, and as noted earlier thresholds for disorder reflect a range of considerations (Kessler *et al.*, 2003).

Second, consider *explanatory pluralism*. Philosophers of science have contrasted a monistic explanatory account which attempts to provide one overall unified account of the phenomena of the world, with an account of science which emphasizes the need for multiple

partial models. The model or metaphor of maps may be used to account for way in which explanatory pluralism itself works; thus, a cartographer may employ multiple different maps of the world, each accounting for different features of reality, and each of which is useful for a particular purpose. As noted earlier, in philosophy of psychiatry there is ongoing debate about the extent to which the biopsychosocial model, which encourages a focus on different dimensions of disease and illness, is merely eclectic, or provides the appropriate scaffolding for considering a range of causal mechanisms.

A major area of debate in philosophy of science regards reductive explanations. It has long been argued that the phenomena of the world can be organized along different levels, ranging from the physical through the biological and on to the social. A reductionist approach aspires to explain higher level theories (e.g., biological models) in terms of lower level accounts (e.g., physical models). Certainly, as science has progressed, such inter-theoretic reduction seems to have occurred; thus we can account for the properties of DNA (which plays such a key role in biology) in terms of its particular structure (that it in terms of its underlying physico-chemical properties) (Bickle, 1998; Brigandt and Love, 2017).

Pluralists have emphasized, however, that such successes are only part of the story of science. Science is often concerned, for example, with phenomena that emerge only at higher levels of organization: these require models that cannot simply be reduced to lower level accounts (Bedau and Humphreys, 2008; Corradini and O'Connor, 2010; De Vreese, Weber and Van Bouwel, 2010). Furthermore, as emphasized in the metaphor of science as cartography, multiple different sorts of models of reality may be useful for different purposes. Focusing on biological science, Sandra Mitchell has concluded, "Given the multiplicity of causal paths and historical contingency of biological phenomena, the type of integration that can occur … will itself be piecemeal and local … pluralism with respect to models can and should coexist with

integration in the generation of explanations of complex and varied biological phenomena" (Mitchell, 2003).

Discussions of pluralism often refer to the relationships between different "levels" of explanation, but "levels" themselves are better understood as ways of referring to different sorts of organizational (part-whole), spatial, and temporal relationships (Eronen, 2021). Depression, for example, occurs on a much slower time scale than the firing of a neuron. Neuroimaging studies occur on a larger spatial scale than does the study of protein synthesis. Slow versus fast and large versus small might carve things up differently than higher versus lower, so that a pluralist approach to explanation is required (Eronen, 2021). In philosophy of science and neuroscience, there is ongoing exploration of how best to conceptualize causal processes and mechanisms, including causality across different levels (Bickle, 2007; Beebee, Hitchcock and Menzies, 2010; Glennan and Illari, 2018; Calzavarini and Viola, 2021). For psychiatry, however, it is key to be aware of the complexity of psychobiological systems, and to avoid overly simplistic neuro-reductionism (Fulford, Bortolotti and Broome, 2014; Frances, 2016; Kendler, Parnas and Zachar, 2020).

Once again, these philosophical constructs have practical import. Applying scientific pluralism to psychiatry, Kendler has argued that first-person subjective experiences and sociocultural factors play a vital role in the etiology of psychiatric disorders, such that this etiological role cannot be captured by focusing on the basic biology of the brain. "A bottom-up hard reductionist approach to psychiatric illness will be futile if basic neurobiological risk factors are frequently modified by higher-order processes, including environmental, psychological, and cultural experiences" (Kendler, 2005). Kendler suggests that a pluralistic psychiatry should aim for "patchy reductionism" and "piecemeal integration" as it tries to

understand the multi-level causal interactions that give rise to psychopathology (Kendler, 2005).

When we think about psychotropic medications, for example, we often think of specific receptor effects. While important, this downplays how these agents exert a cascade of effects, impacting neural networks and ultimately behavior. A pluralistic clinical psychopharmacology is needed in order to flesh out these higher-level mechanisms in greater detail. Further, complex multilevel explanations involving a range of mediating processes are needed to explain higher-level phenomena such as placebo and nocebo effects, and to account for molecular-social Interactions such as how antidepressants acting on serotonergic pathways impact social hierarchy. While the focus of much psychopharmacology has been on lower-level mechanisms, such as receptor actions, a pluralistic approach emphasizes that cognitive and phenomenological processes can also be important psychopharmacological targets (Aftab and Stein, 2022). Analogously, a pluralistic approach may be useful in exploring the causal powers relevant to psychotherapy (Oddli *et al.*, 2022).

Third, consider *value pluralism*. Value pluralism, which emphasizes that there are many different moral values, is typically considered as a position in moral philosophy. However, value pluralism is also relevant to science in general, and psychiatry in particular, in a number of ways. In particular, as we've discussed, choices about how best to classify and describe the structures and mechanisms of the world reflect a range of scientific purposes or epistemic values, and indeed debates about scientific pluralism intersect with debates about science and society (Ludwig and Ruphy, 2021). Differences between DSM-5 and ICD-11, for example, do not necessarily reflect scientific disagreement, but rather acknowledgment of differences in their most important aims and associated values (Stein and Reed, 2019).

Additionally, the argument that natural kinds reflect clusters of properties has been extended to value-laden constructs; "healthiness" for example, may reflect a range of related features, presumably underpinned by relevant biological processes (Boyd, 1988). That said, if we think back to normativist positions on the definition of mental disorder, which emphasize the influence of social and cultural values, different societies and cultures may have different understandings of mental disorder/illness because they value different conceptions of human flourishing (Glackin, 2016).

Philosophical work on value pluralism has long emphasized that given the plurality of values, choices between them will be complex. Early on the philosopher Isaiah Berlin emphasized that values may be incompatible, and this seems consistent with our experience of moral decision-making (Berlin, 1991). Nevertheless, this does not necessarily mean that value-laden choices cannot be made in a reasonable way. Even earlier, Aristotle emphasized the importance of practical wisdom, arguing that a virtuous person succeeds in making correct choices (Kinesella and Pitman, 2012). While practical wisdom may in part involve the application of general principles, Aristotle emphasized the "priority of the particular" in choosing the correct course of action (Nussbaum, 1986).

Value pluralism again has a number of clinical implications. As noted earlier, the psychiatrist-philosopher Bill Fulford and his colleagues have argued that evidence-based health care needs to be complemented by value-based health care (Stoyanov *et al.*, 2021). A growing literature on shared decision-making similarly highlights the important perspectives of those with lived experience of mental illness (Slade, 2017; Fulford and Handa, 2021). Furthermore, given the complexity of psychiatric phenomena, and the relevance of ontological and explanatory pluralism, it seems clear that psychiatric practice and research ought to involve a broad range of disciplines and methodologies. Finally, value pluralism

emphasizes the importance of a range of epistemic virtues, including epistemic humility, including cultural humility (Tervalon and Murray-García, 1998).

Conceptual Tools for Psychiatric Explanation

Four important and interrelated concepts may be useful for psychiatry explanation: *organizational causality, dynamical constitution, downward causality,* and *dual aspectivity* (Kirchhoff, 2015; Fuchs, 2018; Leuridan and Lodewyckx, 2021; Nielsen, 2023).

Organizational causality or 'formal causality' is the idea that the organization or form of an object can have a causal effect on the world. For example, given the same force, a well-made paper airplane flies further than a scrunched-up ball of paper of the same weight, because of the arrangement and form of its surfaces and the way these interact with the surrounding air.

Dynamical constitution is the related notion that objects and processes at smaller scales of enquiry can interact over time in ways that produce objects, systems, or processes at larger scales of enquiry, and that qualities of the larger object can emerge from the interaction between/organization of the component objects and processes (Kirchhoff, 2015; Leuridan and Lodewyckx, 2021). As an example, water has the qualities of a liquid at room temperature because of interaction of dipole H_2O molecules repelling each other.

Downward causality is the idea that these emerging objects, systems, and processes at larger scales of inquiry can entrain, constrain, or otherwise have causal influence over objects at smaller scales. For example, moving water freezes at a lower ambient temperature because of the way that the wider fluid dynamics maintain movement between the molecules (i.e., heat).

Dual aspectivity refers to the idea that whenever talking about a living system, there are at least two perspectives one can take, firstly a body-as-object, naturalistic, or third-person perspective, and secondly a body-as-subject, personalistic, or first-person perspective. Both perspectives are seen to consider the same physical object, but they capture different aspects of the living system/person under study, in line with a pluralist approach (Gallagher, 2017; Fuchs, 2018).

Taken together these concepts provide an approach to understanding constitutionally complex systems such as life forms. Nielsen considers the analogy of a tower of Lego blocks (Nielsen, 2023). During the time a set of Lego blocks constitute a tower, the blocks *are* the tower, but they still exist. Analyses of blocks, their interactions, the structure of the tower, and the surrounding context will all be useful to explanation why the tower falls or stands. Furthermore, the form of the tower is not a separate 'product' of the blocks' organization, nor an epiphenomenal apparition. De Haan puts forward the analogy of a cake: it has clear ingredients, but once baked it has a range of new properties (de Haan, 2020).

In a similar way, organisms are made up of many parts (organs, cells, receptors etc.) and derive properties, such as mindedness, from the complex interactions between these parts in context. Both the parts and the wider organism are no less real because of the knowledge we gain about their parts and how they manage to dynamically constitute a minded creature, and analysis at multiple scales of enquiry—consistent with an emphasis on a complex systems framework and a pluralist approach—is going to be useful for understanding how this creature functions and how things may go awry.

From a clinical perspective, a key lesson that emerges is an appreciation of the complexity involved in clinical formulation and intervention. Given dynamical constitution, neurobiological mechanisms are key to shaping behavior, thoughts, and emotion. However,

given organizational and downward causality, such mechanisms cannot be assumed to have causal primacy within our explanations. Contra the view of the neo-Kraepelinians and overly reductionistic translational neuroscience, mental disorders are not merely brain disorders (Jefferson, 2022). Conversely, interventions such as psychotherapy, may impact brain and body (Brooks and Stein, 2015). Given the complexity of the brain-mind it is unsurprising different kinds of effective psychotherapies have arisen; exploring overlaps and differences, and developing pluralistic or integrative approaches is a useful goal (Cooper and McLeod, 2007; Castonguay *et al.*, 2015).

The biopsychosocial model remains the most influential approach for ensuring a pluralistic approach to assessment and intervention for mental disorders in the clinical context. Despite the criticisms noted earlier, the biopsychosocial model arguably reminds us that organic versus functional distinction is an overly simplistic approach to psychiatric etiology (Spitzer *et al.*, 1992), and that a pluralistic framework which considers a broad range of "difference-makers" is needed in clinical research and practice (Kendler, 2012; Fulford, Bortolotti and Broome, 2014). Tinbergen's model of different sorts of explanations is grounded in evolutionary theory and provides a potentially useful way of ensuring that a range of explanations are included in any clinical formulation (Table 2) (Nesse, 2019).

Jaspers' insistence that both explanations of underlying mechanisms as well as understanding of individual narratives and meanings are important for a comprehensive count remains relevant to contemporary clinical practice. Medical anthropologists have usefully distinguished between disease as a biomedical condition, and illness as the subjective experience of those suffering from such conditions (Kleinman, 1988). Relatedly, work on what has been termed neurophenomenology attempts to integrate neuroscientific knowledge with individual experience (Varela, Thompson and Rosch, 2016; Gallagher, 2017). Finally,

"explanation-aided understanding" – the idea that knowledge of mechanisms can enhance our appreciation of the first person experience - is also a key consideration for improving clinical practice (Kendler and Campbell, 2014).

EMBODIED COGNITION AS A PLAUSIBLE INTEGRATIVE APPROACH

The "mind-body problem" is a paradigmatic issue at the intersection of psychiatry and philosophy. The philosopher Rene Descartes is often cited for his substance *dualism* – that mind and body exist as radically different kinds of substances – and clinicians are typically encouraged to avoid this position in light of a modern naturalist or scientific understanding. At the same time, in their training clinicians are generally not encouraged to explore recent developments in an area of such philosophical complexity, and as a result some implausible assumptions can arise (O'Leary, 2021).

One commonly assumed view is to see the mind as a powerless or 'supervenient' side effect of the physical processes of the brain. Such a view can support neurocentric assumptions, for example, within a biological psychiatry that contends that the brain is where we need to focus the vast majority of our explanatory and treatment efforts. As noted earlier, while the brain is clearly important for understanding mental functioning and mental health, such an approach may be criticized for its neuro-reductionism, where minimal space is made for similarly important aspects of human functioning such as experience, meaning, culture, and context.

Another common view, inspired by the development of computers, is to see the mind as 'software' running on the 'hardware' of the brain. Under such a *computationalist* and *functionalist* view, cognitive functioning is understood as a form of information processing

where the brain takes sensory input and computes appropriate responses. Such an assumption can be seen in the notions of cognitive biases and core beliefs within CBT, with these concepts effectively performing the role of 'bugs in the software' altering our perception of the veridical world. While this can be a useful metaphor, there are multiple issues with this perspective. It is difficult to see how such a view can be integrated with a biological perspective, in which neurons and behaviors are complexly intertwined. Indeed, such a view seems implausible; living creatures are not computers with set functions and this analogy may limit our insight.

Moving beyond assumptions of supervenience and computationalism, *embodied cognition* represents a biologically plausible and strongly integrative view of the mind-body relationship, whereby factors across the biopsychosocial spectrum are considered to have potential explanatory value (Aftab and Nielsen, 2021). Such an embodied perspective has gained momentum within philosophy of psychiatry in recent years; e.g., (Barrett et al., 2016; Bruineberg & Rietveld, 2014; Freund et al., 2016; Tschacher et al., 2017) but is not yet broadly recognized by clinicians nor discussed in training programs. Engagement with embodied understandings of the connection between mind and body is a third key development in philosophy of psychiatry.

Applied to psychopathology, notions of *embodiment*, alongside related ideas such as *embedment, extension,* and *enactivism,* which we will soon unpack, represent one plausible integrative frame for the study and treatment of mental disorder. We will argue that an embodied approach has the potential to incorporate and build on many of the recent conceptual developments highlighted in previous sections, while also cohering well with other contemporary theoretical and methodological developments in a range of disciplines. In this section, we will first define some key terms and review the development of embodied

cognition. We then go on to discuss the application of this approach to the study and treatment of mental disorder.

What is Embodied Cognition?

Embodied cognition refers to a diverse range of approaches across multiple disciplines within cognitive science, including but not limited to psychology, neuroscience, philosophy, robotics, and artificial intelligence. Embodied cognitive science is united by a common interest in moving away from a 'cognitivist' or 'computationalist' view, where the brain is seen as an isolated 'seat of cognition' that receives sensory information, represents the world, and computes appropriate responses to it. Instead, embodied approaches variously emphasize the role of the body and context both in the moment-to-moment constitution of cognition and in the shaping of cognition across development, thus decentering the brain and ideas of representation in how we seek to understand the mind (Shapiro, 2019). Instead of understanding the mind through implicit analogy to computers, embodied approaches seek to understand the mind through analogy to complex living systems adapting to the dynamics of their environment.

Historically the development of embodied cognition has many roots. The most commonly recognized of these roots include: a rejection of a traditional symbolic-representational view of cognition where the experienced world is a model/representation of reality; an interest in expanding upon the success of minimally representational 'connectionist' understandings of cognition such as exampled by neural networks; the emphasis of pragmatic philosophers such as John Dewey on how knowledge entails interaction with the world; phenomenological insights by authors such as Merleau-Ponty that

the body is an intrinsic part of our experience-of and engagement-with the world; work in developmental psychology by Jean Piaget and others who have emphasized interaction with the world over time; and inspiration from the success of dynamic systems theory in modelling the behavior of complex systems (Newen, De Bruin and Gallagher, 2018). Such historical antecedents have converged to produce understandings of the mind that tend to be more plausibly 'organic' and that recognize a broad range of influences shaping human cognition, from genes and molecules to culture and context.

Embodied cognitive science is a diverse field. This is true to the point that the very word 'embodied' can take on subtly different and overlapping meanings in different contexts. It is therefore important to specify the sense in which we use this term. In a summary review of embodied cognition, Shapiro and Spaulding highlight three different yet overlapping themes within the various usages of the term 'embodied'. They refer to these three different themes of overlapping meaning as 'constitution', 'conceptualization', and 'replacement' (Shapiro and Spaulding, 2021). In this article, we generally refer to the *constitutional* understanding of embodiment. It is however worth briefly explaining all three senses of the term as discussed by Shapiro and Spaulding.

In the *replacement* sense of 'embodied', emphasis is on the need to replace our systems of understanding the mind with less representational and more dynamical ones. In other words, developing ways of understanding the brain, not as generating a mirror-like re-presentation of the world, but rather as resonating with the world directly. A classic example would concern how best to think about the action of catching a fly ball in baseball. Rather than representing the entire environment and computing the ball's trajectory, a non-representational and embodied view would suggest that we engage with simple visual strategies in order to ensure that we are in an optimal position to catch it (Wilson and

Golonka, 2013). This sense of embodiment is particularly associated with the position of Radically Enactive Cognition which attempts to understand cognitive processes with no refence to representation (Hutto and Myin, 2013).

When the term 'embodied' is used in a *conceptualization* sense, the focus is on psychological concepts and processes, and how they are shaped by the kinds of bodies and experiences we have. The key idea is that the way we conceptualize the world would likely be different if we had different sorts of bodies to navigate with. For example, consider the idea that we think in terms of 'up' as metaphorically connected to positivity and action and 'down' as connected to depression and inaction not simply as a cultural quirk but because of shared associations rooted in our bodily experiences and actions (Lakoff and Johnson, 1999). Accumulating behavioral and neurobiological evidence supports the related ideas that 1) there is significant overlap in the neural processes involved in sensorimotor co-ordination and those involved in so-called 'higher' cognitive and social processing, and 2) such overlap means that 'higher' cognitive processes are not siloed in the brain – but are influenced by bodily and sensorimotor context such as posture, current action, and internal physiological state (Gallese and Lakoff, 2005; Seth and Friston, 2016).

When used in a *constitutional* sense, which is our main focus here, 'embodied' refers to the idea that mental processes are best thought of as not constituted by the brain alone, but rather as emerging from the brain and body acting in concert, i.e., as one extended system. The mind does not arise from the efforts of the brain to represent the world, but rather is an active process of the entire organism navigating, adapting to, and making sense of the world (Lakoff and Johnson, 1999; Thompson, 2010; Shook and Solymosi, 2014; Engel, Friston and Kragic, 2015; Varela, Thompson and Rosch, 2016; Johnson, 2017). On such a view, for example, the release of cortisol and adrenaline from the adrenal glands in response to an

acute stressor is not simply an event occurring on one level of analysis with a modulating effect on cognition at a higher level of analysis, but rather is *part of* a single, body-involving, cognitive-affective response to threat. Thus the processes that constitute emotions weave in and out of the brain, and include a range of interoceptive components (Seth and Friston, 2016; Colombetti, 2017).

Now that we have outlined what is meant by 'embodiment' it will be useful to define some related ideas, specifically those of *embedment*, *extension*, and *enactivism*. Together these ideas, alongside embodiment, are often discussed under the umbrella term '4E'. Sometimes '5E' is also used, typically in reference to a focus on emotion and affectivity.

Embedment is the idea that cognitive functioning involves constant casual interplay with the environment across multiple timescales. Consequently, in order to understand cognition, recognition of the role of context is vital. When considering human functioning, the environment is also considered not simply to be a physical one, but a social-cultural one, constituted by others, alongside their artifacts and shared structures of meaning (Durt, Fuchs and Tewes, 2017). Embedment highlights that across the timescales of evolutionary change, socio-cultural development, life-span learning, and moment to moment cognition, human beings are both deeply influenced by, and in turn influence, their surrounding environments. Even the most abstract and intellectual activities, such as mathematics, entail a thoroughly embodied and embedded skill set (Lakoff and Johnson, 2010; Gallagher, 2017).

Extension is an idea that is in many ways similar to embedment, but makes a more radical claim. Specifically, extension refers to the idea that cognitive process are often best understood as extending out beyond the body and looping through the world (Clark and Chalmers, 1998). To continue the mathematics example, rather than merely understanding a calculator as supporting the cognitive processes of an individual, an extended view of mind

would hold that the calculator becomes part of the cognitive process. In a well-known thought-experiment, Clark and Chalmers contrast Inga—who navigates from memory, with Otto—who has Alzheimer's disease and relies on written directions in a notebook. Given that the only difference between the two cases is that the process of navigation takes place wholly inside the brain in Inga's case and partly outside it in Otto's, they argue that it is arbitrary to confine cognition to what occurs within the confines of the skull (Clark and Chalmers, 1998).

Enactivism is an idea that subsumes and builds upon ideas of embodiment and embedment. It may be explained in different ways and with different points of emphasis, but here we focus on *autopoietic enactivism* (Ward, Silverman and Villalobos, 2017). 'Autopoietic' simply means 'self-creating'. Accordingly, within autopoietic enactivism, the focus is on the idea that mindedness is brought forth, or rather enacted, through the organizational structure of life forms and their efforts to constantly maintain themselves within the context of their environment. An enactive perspective holds that life forms are shaped through evolution to try and survive, and that this inherent purpose sets up the necessary groundwork for the emergence of relational *meaning,* i.e., meaning for an organism. In order to survive, organisms have to evolve to, or learn how to seek food, avoid predators, and so on – that is, to respond differentially to affordances in the world (Dotov, Nie and de Wit, 2012). Cognition is *sense-making* – a constantly unfolding process, one that is body and action involving, relational, and inherently affective/meaning-laden. Thus, cognition is not a linear process of sensation-perception-cognition-action, but rather a circular process of sensory-motor engagement. And, rather than a neurocentric view in which brain takes in information and represents a model of the world, the brain is instead an organ of co-ordination, learning, and mediation within this sensorimotor loop (Fuchs, 2018; Gallagher, 2018).

A prescient development in theoretical neurobiology is an enactive formulation of sentient behaviour and situated cognition called *active inference* (Friston et al., 2011; Parr et al., 2018). Active inference underwrites one cornerstone of computational psychiatry (Montague et al., 2012); namely, the use of computational neuroscience and machine learning to understand psychiatric conditions. In part, active inference dissolves the dialectic between enactivism and computationalism by noting that any embodied agent can be described 'as if' they were inferring the causes of their sensorium; crucially, these causes include *their own actions on the world*. This has the advantage of being able to formulate choices and behaviour in terms of (subpersonal, Bayesian) belief updating and notions such as *planning as inference* (Botvinick & Toussaint, 2012). In other words, it offers the prospect of naturalising psychology as inference and psychopathology as false inference (e.g., inferring something is there, when it is not; such as in hallucinations and delusions. Conversely, inferring something is not there, when it is; such as neglect and dissociation syndromes). Much of the focus—in this kind of computational psychiatry—is on decision-making and covert (mental) action, read as attention (Parr & Friston, 2017). Computationally, this mental action is often cast as affording precision (i.e., increasing postsynaptic gain) to sources of sensory evidence and (prior) beliefs. This provides a link between psychopathology and neuromodulation on the one hand (Friston, 2022), while foregrounding the importance of interactions with the lived world on the other. We will return to active inference, in the setting of communication and (cultural) niche construction, later.

Overall, the embodied/4E understanding of cognition presents a non-dualistic understanding of the mind that appears biologically plausible and does not fall prey to reductionistic temptation. Human functioning is understood from this perspective in a way that preserves a sense of agency, while also recognizing the diverse array of influences that

shape and influence human health and behavior, from genes to culture. It is therefore an integrative view, demanding both current and historical analysis of the entire brain-body-environment system if we are to understand patterns in human behavior and cognition. This approach is consistent with a coordinated pluralism (Sullivan, 2017), and arguably even with an integrative pluralism (Gauld *et al.*, 2022), and has led to the suggestion that accounts of mental disorder grounded in embodied/4E cognition may represent a path to solving the 'integration problem' in psychiatry; i.e., the fact that we have identified causal factors across the brain, body, and environment, but struggle to conceptualize how these causes come together to shape patterns of disorder (de Haan, 2020).

Clinical Applications of Embodied/4E Cognition

Recently several conceptual frameworks grounded in the embodied/4E perspective have been applied to mental disorders as a whole (Maiese, 2016; Fuchs, 2018; de Haan, 2020; Nielsen, 2023). These frameworks view mental disorders as representing *disruptions to sense-making*, a view that is consistent with attempts to bridge the naturalist-normativist divide (Maiese, 2021). They also share a vision that embodied/4E cognition serves as an integrative framework for the conceptualization, study, and treatment of these conditions. Additionally, there have been several efforts to develop descriptive and explanatory models of particular mental disorders from an embodied/4E perspective (de Haan *et al.*, 2013; Fuchs, 2013; Tschacher, Giersch and Friston, 2017; Ramírez-Vizcaya and Froese, 2019; Glackin, Roberts and Krueger, 2021; Gallagher, 2022).

A focus on embodied cognition leads to a view of mental disorders as *constitutionally complex*, involving biological, cognitive-emotional, environmental, and socio-cultural aspects.

This perspective emphasizes both *biology and agency,* acknowledging biological scales of enquiry as relevant without reducing the explanatory importance of experience and choice. It also incorporates ideas of organizational causality, downward causality, and dynamical constitution to break down the received mind-body divide, and aligns well with the notion of mental disorders as fuzzy mechanistic property clusters (Kendler, Zachar and Craver, 2011).

Through the notion of embedment, these frameworks emphasize the active and historical role of the physical and sociocultural environment. All organisms, particularly humans, are deeply historical and ecologically informed creatures. Shaped by our evolutionary, socio-cultural, and developmental pasts, we are understood to strive to adapt to the present context and predicted future (Thompson, 2010; Durt, Fuchs and Tewes, 2017) Applied to psychiatry this allows integration with perspectives such as evolutionary psychiatry (Nesse, 2019), cultural psychiatry (Kleinman, 1988; Kirmayer and Ramstead, 2017), and developmental psychopathology (Cicchetti, 2016). Applications of computational psychiatry — in particular, active inference — have embraced this view, providing a pluralistic, yet formal and mechanistic, account of conditions such as depression and obsessional compulsive disorder; often with a special focus on interoception and bodily states (Constant et al., 2021; Duquette, 2017; Kiverstein et al., 2019; Rae et al., 2019).

In the embedded view, however, culture is not seen only as a historical force having influence across development, but also as a living context. In this 'constitutional view of culture', culture is seen as a 'shared world' or structure of knowledge, meaning, and artifact, constituted by ongoing engagement (Durt, Fuchs and Tewes, 2017). Such a shared world represents a historical context for the development of an individual and the way they make sense of the world, but also continues to play out in the moment-to-moment interaction of individuals, including in the clinical encounter. Embeddedness therefore pushes clinicians to

actively consider the role of culture both in the lives and histories of their patients, but also in the clinician-patient interaction.

Via the notion of enactivism, these frameworks subscribe to a *process orientation,* with mental disorders not viewed as static problems/dysfunctions in the brain or psyche, but rather as constantly unfolding patterns of how we make sense of and engage with the world. Through interactions with their specific environment and its particular affordances, thinking beings create and discover meaning for themselves. Rather than stemming from some underlying 'cognitive error' or 'psychic disturbance', mental disorders emerge within the circular relationships between patient and world – as a maladaptive pattern of sense-making (Nielsen, 2023).

This process-orientation affords accords well with the focus of neuroscience and computational psychiatry on *active inference*, whereby predictive processing frameworks formally model how organisms develop probabilistic assessments of their environment so as to adapt optimally. Indeed, several authors have considered how best to integrate such frameworks with embodied/4E approaches, noting that the brain-mind, including interoceptive components, engage in embodied predictive processing in order to maintain enactive engagement with the environment (Seth and Friston, 2016; Allen and Friston, 2018; Gallagher and Allen, 2018). In their embodied/4E account, Friston and colleagues suggest the term *enactive inference* (Ramstead, Kirchhoff and Friston, 2020).

In the clinical setting, given the thoroughgoing role of affordances and affectivity within the enactive view, a process-orientation accords with psychotherapies that draw patients' attention to early maladaptive schemas and current emotional dynamics in order to better learn to navigate them (Riso *et al.*, 2007; Greenberg, 2012) – an exercise in sense-making about sense-making (Nielsen, 2023). Further, in the enactive perspective, therapeutic

interventions in psychiatry improve the fit between the individual and their environment. This can in turn be achieved either by altering the sense-making and behavior of the individual, or by changing the world around them. This entails integration with notions of social psychiatry, as well as environmentally focused mental health interventions.

Consider for example, an embodied/4E approach to addiction (Glackin, Roberts and Krueger, 2021). Addiction is embodied insofar as the impact of substances on neurobiological mechanisms govern mental activity, it is embedded insofar as such activity unfolds within and depend upon the environment, and it is enacted insofar as the addict experiences environmental affordances in particular ways due to addiction-related desires. In this framework, a merely neural basis of addiction is replaced with a view of addiction as simultaneously neuronally-and-externally constituted, and there is an account of how drug taking transforms the addict's world – altering their personal agency and lived experience.

There are several advantages of such a view. First, this perspective allows an integration of neurobiological accounts of addiction with accounts of the lived experience of the addict. Second, the binary choice of seeing addiction as a medical disease versus a personal choice — a key issue in the philosophy of addiction (Pickard and Ahmed, 2019; Glackin, 2020) — can be seen to be an overly simplistic false dichotomy. Third, this perspective enables us to reconsider strategies for recovery; in particular, it provides an account of how the addict may be able to change their lived experience by manipulating the environment and altering its affordances, so that there is a change in the dynamic interaction of brain biology, interoception, and surrounding context.

Another instructive example is the work of Tschacher and colleagues on schizophrenia (Tschacher, Giersch and Friston, 2017). These authors point out that sensorimotor dysfunctions are closely associated with psychotic symptoms, leading to altered timing in the

processing of stimuli and to disordered appraisals of the environment. They argue, therefore, that problems of social cognition can be viewed as disordered embodied communication. Finally, they suggest that this account suggests novel treatment strategies through body-oriented interventions. Again, then an embodied/4E approach is theoretically able to integrate biological and phenomenological perspectives, and also has practical implications for the clinical context.

.

DISCUSSION

Psychiatry is an inherently conceptually laden and conceptually complex field. Yet the opportunity to reflect on the concepts implicit in psychiatric practice arises infrequently for both clinicians and trainees. Instead, a range of tacit assumptions may be invoked - about the nature of mental disorders, of diagnostic categories, of causal explanations, and the mind - that inform daily engagements with psychiatric taxonomy, clinical assessment and diagnosis, and the discussion of conditions and treatments with patients. These tacit assumptions have, however, been carefully addressed by philosophy of psychiatry, and here we have reviewed key advances in this field, and their clinical implications.

First, in conceptualizing and categorizing mental disorder, both naturalist and normativist considerations have emerged as important – the field increasingly accepts that such work entails a consideration of both facts and values. At a conceptual level this encourages moving away from strict scientism to soft naturalism – a position that embraces both psychobiological mechanisms and personal agency. In clinical practice the bridging of naturalism and normativism facilitates moving away from an approach in which disorders are reified, and toward appropriately comprehensive and individualized evaluations of patients.

Awareness of the importance of facts as well as values may also facilitate both evidence-based and values-based mental health care.

Second, in considering the nature of psychiatric science, there is now increasing emphasis on a pluralist approach, including ontological, explanatory, and value pluralism. Conceptually, this approach again shifts away from a reductionistic scientism towards a liberal naturalism, and is consonant with Jaspers' early pluralist approach, encompassing both explanatory accounts of behavior as well as an understanding of the individual person – an approach which has since been espoused by a wide range of philosophers (De Caro and Macarthur, 2022). Clinically this view emphasizes the importance of a broad range of causal "difference-makers" as well as considerations of "lived experience" in both research and practice.

Third, in considering a range of questions about the brain-mind, and how both somatic and psychic factors are key in mental disorders, conceptual and empirical work on embodied cognition provides an increasingly valuable approach. Viewing the brain-mind as embodied, embedded, and enactive provides further support for a conceptual approach to the mind-body problem that facilitates the integration of advances in both cognitive-affective neuroscience and phenomenological psychopathology, as well as in a range of other disciplines. Work on embodied cognition is gaining increasing purchase in rethinking mental health and mental disorders (Maiese, 2016; Fuchs, 2018; de Haan, 2020; Nielsen, 2023).

Conceptual competence has various elements, including making explicit conceptual assumptions, developing a philosophical vocabulary, acquiring familiarity with relevant frameworks, and maintaining a degree of "conceptual humility" (Aftab and Waterman, 2021). Several of these elements have been exemplified in the current paper, but at this point we would wish to emphasize the virtue of epistemic humility in particular. Philosophy of nosology

has taught us that despite the enormous amount of work that has been done to improve our classifications and criteria, these remain tentative, and reification of putative entities must be strenuously avoided.

Similar considerations would apply to our attempts here to advance the philosophy of psychiatry. We began by noting that progress in philosophy has been disputed, and we are wary of claiming too much. Some of the issues in philosophy of psychiatry date back to ancient times, and some of the constructs employed seem to be "essentially contested": there are a range of competing views, and there deserves to be ongoing discussion and debate (Gallie, 1955). This phrase may well apply to the concept of disease, which may be intractably messy (Bortolotti, 2020; Kukla, 2022). Still, the purpose of philosophy is not necessarily to resolve every dispute or to eradicate disagreement, but rather to properly articulate and understand them. To the extent that the issues considered here have been more rigorously articulated in theoretical work, and more thoughtfully considered in the clinical context, some progress may be claimed.

Several limitations of the approach taken here deserve particular emphasis. First, we have summarized arguments and conclusions from the literature, rather than attempting to rigorously defend any particular stance. Relatedly, we have not had space to address ongoing work and key variants of some the positions that have been put forward here, nor important critiques of these positions (Wakefield, 2005; Parfit, 2017; Nielsen, 2021; De Caro and Macarthur, 2022; Russell, 2023). Key constructs employed here – including soft naturalism, pluralism, and embodied cognition – all deserve much deeper consideration.

Second, while have drawn some links between these particular constructs – soft naturalism, pluralism, and embodied cognition – our view is that much further work along these lines is possible. At the broadest level, some work on these constructs seems to allow

for a degree of rapprochement between analytic and continental philosophy (Egginton and Sandbothe, 2004), while at a more localized level, it seems to us that linkages can be made between key philosophers who have spoken to these issues (including John Dewey – who prefigured notions of embodied cognition, Wilfred Sellars – who contrasted the scientific and manifest image, and Hilary Putnam – who contributed to work on the fact/value distinction).

Third, we have been selective in our focus on progress in the field, omitting large swaths of work in philosophy of psychiatry – not the least being ethics – and potentially downplaying a range of authors and advances. The breadth and depth of work by both ancient and modern philosophers who have considered the questions raised here is extraordinary, and we would encourage readers to explore further (Fulford, Thornton and Graham, 2006; Sadler, van Staden and Fulford, 2015; Aftab and Waterman, 2021; Tekin and Bluhm, 2022).

Given our interests in philosophy, psychiatry, and the cognitive-affective sciences, we cannot but also ask the extent to which there has been work on the cognitive science of philosophy and psychiatry themselves. In this review, for example, we have focused on developments in embodied cognition. But an embodied cognition perspective suggests, for instance, that when we think about categories such as mental disorders, rather than being aware how deeply reliant we are on embodied metaphors, we are instead prone to essentialize them (Adriaens and De Block, 2013; Stein, 2013). We may well have particular difficulty in getting our heads around relevant constructs such as the bridging of naturalism and normativism, the avoidance of essentialism and reification, and the possibility of "responsibility without blame".

Notably, a number of key themes have emerged from the different parts of this paper; we will highlight three. First, a key theme has been that of integration. Thus, we have discussed the integration of elements of naturalism and normativism, of evidenced-based

care with values-based care, of knowledge of science with the understanding of experience, and of psychobiological mechanisms with personal agency. Further, we have emphasized the potential value of the embodied/4E approach in integrating a range of disciplines concerned with the brain-mind, including cognitive-affective neuroscience, developmental psychopathology, and social psychiatry.

A related theme has been that of "balance". Our notions of mental disorders need to avoid the poles of scientism and relativism, our explanations need to avoid both neuro-reductionism and culturalism, and our approach to the mind-body problem should be one that avoids both a brainless and a mindless psychiatry (Lipowsky, 1989). Our introduction mentioned Aristotle's notion of the golden mean, and the emphasis in Chinese philosophy on the balance of yin and yang; a balanced perspective that is able to judiciously weigh up a range of principles and particulars surely lies at the heart of good clinical work (Radden and Sadler, 2010).

A final theme has been that of complexity. We have argued that it is important to avoid essentialism and reductionism in psychiatry, and that clinical assessments need to go far beyond our diagnostic criteria to assessing a range of domains, and to understand each patient as a unique individual. While monocausal models of disease entities (e.g. Treponema pallidum causing neurosyphilis) have been very useful, contemporary psychiatry requires a coordinated (Sullivan, 2017) or integrative (Gauld *et al.*, 2022) pluralism. And the embodied/4E perspective emphasizes the complexity of the living being's dynamic engagement with their environments over time. The complexity of the brain-mind and of clinical conditions is a key reason that calls for simplistic paradigm shifts in psychiatry are unlikely to succeed (Stein *et al.*, 2022); instead – in keeping with philosophy as the love of wisdom – there is a need for a 'wise psychiatry'.

The philosophical resources here may be useful in considering the extent to which psychiatry has made progress in the past, our current balance of success and failure, and our future aspirations. It seems to us incontrovertible that philosophy has played a key role in psychiatry, whether implicitly or explicitly, and that it will continue to do so into the future. Our immediate hope is that work in philosophy of psychiatry will contribute to the conceptual competence of clinicians; by rendering the implicit explicit, philosophical analysis may help expose the implications, limitations, contradictions and even absurdities that potentially underlie received ideas and prominent positions. Our long-term hope is that advances in philosophy of psychiatry will in turn have positive clinical impact, contributing to integration, balance, and wisdom in psychiatric practice.

Table 1: The Naturalist-Normativist Spectrum

This table outlines the various levels, steps, or questions in the naturalist-normativist debate. Each question is followed by a list of possible responses.

Are there biological and behavioral states that can be characterized as dysfunctional or malfunctional in objective terms independent of human interests?

- Yes, dysfunction can be described in entirely value-free terms.
- Dysfunction may not be explicitly defined in value-laden terms, but an evaluative component or human interests will play a role when the concept is operationalized in a particular context.
- No. There are biological and behavioral processes but the characterization of these processes as 'dysfunctional' is not an objective fact independent of human interests.

Is there an essence that is shared by all dysfunctions?

- Yes. The malfunctions are grouped together because they share an essence (e.g., they are all failures of a mechanism to perform a function for which it was naturally selected).
- No. Dysfunction refers to a family of related concepts, such that there is no one account of it that is uniquely correct or uniquely privileged, and there is no common characteristic that is shared by all. Some of these concepts may be value-free and others may be value-laden.
- Yes. When these processes have in common is a particular social or folk-psychological judgement of abnormality.

Is "dysfunction" necessary for disorder status?

- Yes
    - Necessary and sufficient
    - Necessary but not sufficient – a harm component or a human-interest component is also necessary
- No. Disorder judgments may be legitimately made in the absence of explicit "dysfunction" judgment (e.g., based on considerations of biological or behavioral regularities – "mechanistic property clusters" – and harm)

What are the relevant human interests?

- Diverse considerations of harm (distress, disability, risk, etc.)
- Diverse clinical and scientific interests that arise in different contexts
- Diverse stakeholders' interests and values
- Sociocultural norms (social deviance)
- Functional norms of self-maintenance and adaptation

Table 2: Tinbergen's 4 Causal Mechanisms in the Clinic (after Nesse, 2019)

		Two Objects of Explanation	
Two Kinds of Explanation		**Diachronic/ developmental**	**Synchronic/ current trait**
	Proximate	**Ontogeny** How did the relevant mechanisms develop in this individual?	**Mechanism** What structures, processes, and mechanisms underly the particular trait?
	Evolutionary	**Phylogeny** How did the relevant mechanisms develop in this species?	**Adaptation** What is the adaptive significance of the particular trait?

REFERENCES

Adriaens, P.R. and De Block, A. (2013) 'Why We Essentialize Mental Disorders', *Journal of Medicine and Philosophy*, p. jht008. Available at: https://doi.org/10.1093/jmp/jht008.

Aftab, A. and Nielsen, K. (2021) 'From Engel to Enactivism: Contextualizing the Biopsychosocial Model', *European journal of analytic philosophy*, 17(2), pp. 5–22. Available at: https://doi.org/10.31820/ejap.17.2.3.

Aftab, A. and Ryznar, E. (2021) 'Conceptual and historical evolution of psychiatric nosology', *International Review of Psychiatry*, 33(5), pp. 486–499. Available at: https://doi.org/10.1080/09540261.2020.1828306.

Aftab, A. and Stein, D.J. (2022) 'Psychopharmacology and Explanatory Pluralism', *JAMA Psychiatry*, 79(6), p. 522. Available at: https://doi.org/10.1001/jamapsychiatry.2022.0470.

Aftab, A. and Waterman, G.S. (2021) 'Conceptual Competence in Psychiatry: Recommendations for Education and Training', *Academic Psychiatry*, 45(2), pp. 203–209. Available at: https://doi.org/10.1007/s40596-020-01183-3.

Allen, M. and Friston, K.J. (2018) 'From cognitivism to autopoiesis: towards a computational framework for the embodied mind', *Synthese*, 195(6), pp. 2459–2482. Available at: https://doi.org/10.1007/s11229-016-1288-5.

Amoretti, M.C. and Lalumera, E. (2022) 'Wherein is the concept of disease normative? From weak normativity to value-conscious naturalism', *Medicine, Health Care and Philosophy*, 25(1), pp. 47–60. Available at: https://doi.org/10.1007/s11019-021-10048-x.

Andreasen, N.C. (1985) *The broken brain: the biological revolution in psychiatry*. New York Cambridge Philadelphia San Francisco London Mexico City São Paulo Singapore Sydney: Harper & Row, Publishers.

Andreassen, O.A. *et al.* (2023) 'New insights from the last decade of research in psychiatric genetics: discoveries, challenges and clinical implications', *World Psychiatry*, 22(1), pp. 4–24. Available at: https://doi.org/10.1002/wps.21034.

Barry, E.S., Merkebu, J. and Varpio, L. (2022) 'State-of-the-art literature review methodology: A six-step approach for knowledge synthesis', *Perspectives on Medical Education*, 11(5), pp. 1–8. Available at: https://doi.org/10.1007/S40037-022-00725-9.

Bayer, R. (1987) *Homosexuality and American Psychiatry: The Politics of Diagnosis*. Princeton University Press.

Bechtel, W. (2007) *Mental Mechanisms*. Routledge.

Bedau, M. and Humphreys, P. (2008) *Emergence: Contemporary Readings in Philosophy and Science*. MIT Press.

Beebee, H., Hitchcock, C. and Menzies, P. (2010) *The Oxford Handbook of Causation*. Oxford University Press.

Beresford, P. and Russo, J. (eds) (2022) *The Routledge international handbook of mad studies*. Milton Park, Abingdon, Oxon ; New York, NY: Routledge (Routledge international handbooks).

Berlin, I. (1991) *The Crooked Timber of Humanity*. Fontana Press.

Bickle, J. (1998) *Psychoneural reduction: the new wave*. Cambridge, Mass: MIT Press.

Bickle, J. (2007) *The Oxford Handbook of Philosophy and Neuroscience*. Oxford University Press.

Blackford, R. and Broderick, D. (eds) (2017) *Philosophy's future: the problem of philosophical progress*. Hoboken: Wiley.

Blashfield, R.K. (1982) 'Feighner et al., Invisible Colleges, and the Matthew Effect', *Schizophrenia Bulletin*, 8(1), pp. 1–6. Available at: https://doi.org/10.1093/schbul/8.1.1.

Bolton, D. and Gillett, G. (2019) *The Biopsychosocial Model of Health and Disease: New Philosophical and Scientific Developments*. 1st ed. 2019. Cham: Springer International Publishing : Imprint: Palgrave Pivot. Available at: https://doi.org/10.1007/978-3-030-11899-0.

Boorse, C. (1977) 'Health as a theoretical concept', *Philosophy of Science*, 44(4), pp. 542–573. Available at: https://doi.org/10.1086/288768.

Boorse, C. (1997) 'A Rebuttal on Health', in J.M. Humber and R.F. Almeder (eds) *What Is Disease?* Totowa, NJ: Humana Press (Biomedical Ethics Reviews), pp. 1–134. Available at: https://doi.org/10.1007/978-1-59259-451-1_1.

Boorse, C. (2014) 'A second rebuttal on health', *Journal of Medicine and Philosophy*, 39, pp. 683–724.

Bortolotti, L. (2020) 'Doctors without "Disorders"', *Aristotelian Society Supplementary Volume*, 94(1), pp. 163–184. Available at: https://doi.org/10.1093/arisup/akaa006.

Boudry, M. and Pigliucci's, M. (2017) *Science Unlimited? The Challenges of Scientism*. University of Chicago Press.

Boyd, R. (1988) 'How to Be a Moral Realist', in G. Sayre-McCord and (ed) (eds) *Essays on Moral Realism*. Cornell University Press.

Boyd, R. (1999) 'Homeostasis, species, and higher taxa', in R. Wilson and (ed) (eds) *Species: New Interdisciplinary Essays*. MIT Press.

Brandenburg, D. (2018) 'The Nurturing Stance: Making Sense of Responsibility without Blame: The Nurturing Stance', *Pacific Philosophical Quarterly*, 99, pp. 5–22. Available at: https://doi.org/10.1111/papq.12210.

Brandenburg, D. and Strijbos, D. (2020) 'Reproach without Blameworthiness', *Philosophy, Psychiatry, & Psychology*, 27(4), pp. 399–401. Available at: https://doi.org/10.1353/ppp.2020.0051.

Brigandt, I. and Love, A. (2017) 'Reductionism in Biology', in *The Stanford Encyclopedia of Philosophy*.

Brooks, S. and Stein, D.J. (2015) 'Psychotherapy and neuroimaging in anxiety and related disorders', *Dialogues in Clinical Neuroscience*, 17, pp. 287–293.

Bueter, A. (2019) 'A Multi-Dimensional Pluralist Response to the DSM-Controversies', *Perspectives on Science*, 27(2), pp. 316–343. Available at: https://doi.org/10.1162/posc_a_00309.

Calzavarini, F. and Viola, M. (eds) (2021) *Neural mechanisms: new challenges in the philosophy of neuroscience*. Cham, Switzerland: Springer (Studies in brain and mind, volume 17).

Campbell, J.K., O'Rourke, M. and Slater, M.H. (eds) (2011) *Carving nature at its joints: natural kinds in metaphysics and science*. Cambridge, Mass: MIT Press (Topics in contemporary philosophy).

Cartwright, N. (1999) *The Dappled World: A Study of the Boundaries of Science*. Cambridge University Press.

Castonguay, L.G. *et al.* (2015) 'Research on psychotherapy integration: Building on the past, looking to the future', *Psychotherapy Research*, 25(3), pp. 365–382. Available at: https://doi.org/10.1080/10503307.2015.1014010.

Chapman, R. (2021) 'Neurodiversity and the Social Ecology of Mental Functions', *Perspectives on Psychological Science*, 16(6), pp. 1360–1372. Available at: https://doi.org/10.1177/1745691620959833.

Chemla, K. and Keller, E. (2017) *Culture without Culturalism: The Making of Scientific Knowledge*. Duke University Press.

Cicchetti, D. (ed.) (2016) *Developmental psychopathology*. Third edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

Clark, A. and Chalmers, D. (1998) 'The extended mind', *Analysis*, 58, pp. 10–23.

Colombetti, G. (2017) *The feeling body: affective science meets the enactive mind*. First MIT Press paperback edition. Cambridge, Massachusetts London, England: The MIT Press.

Conley, B.A. and Glackin, S.N. (2021) 'How to Be a Naturalist and a Social Constructivist about Diseases', *Philosophy of Medicine*, 2(1). Available at: https://doi.org/10.5195/pom.2021.18.

Cooper, M. and McLeod, J. (2007) 'A pluralistic framework for counselling and psychotherapy: Implications for research', *Counselling and Psychotherapy Research*, 7(3), pp. 135–143. Available at: https://doi.org/10.1080/14733140701566282.

Cooper, R. (2004) 'What is wrong with the DSM?', *History of Psychiatry*, 15(1), pp. 5–25. Available at: https://doi.org/10.1177/0957154X04039343.

Corradini, A. and O'Connor, T. (2010) *Emergence in Science and Philosophy*. Routledge.

Craver, C. (2007) *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Clarendon.

Cuthbert, B.N. and Insel, T.R. (2013) 'Toward the future of psychiatric diagnosis: the seven pillars of RDoC', *BMC Medicine*, 11, p. 126.

De Caro, M. and Macarthur, D. (eds) (2022) *The Routledge handbook of liberal naturalism*. Abingdon, Oxon ; New York, NY: Routledge (Routledge handbooks in philosophy).

De Vreese, L., Weber, E. and Van Bouwel, J. (2010) 'Explanatory pluralism in the medical sciences: Theory and practice', *Theoretical Medicine and Bioethics*, 31(5), pp. 371–390. Available at: https://doi.org/10.1007/s11017-010-9156-7.

Dewey, J. (1938) 'Unity of Science as a Social Problem', in *International Encyclopedia of Unified Science*. Chicago: Chicago University Press.

Dings, R. and Glas, G. (2020) 'Self-Management in Psychiatry as Reducing Self-Illness Ambiguity', *Philosophy, Psychiatry, & Psychology*, 27(4), pp. 333–347. Available at: https://doi.org/10.1353/ppp.2020.0043.

Dotov, D.G., Nie, L. and de Wit, M. (2012) 'Understanding affordances: history and contemporary development of Gibson's central concept', *Avant*, 3, pp. 28–39.

Dupré, J. (1993) *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Harvard University Press.

Dupré, J. (2012) *Processes of Life: Essays in the Philosophy of Biology*. Oxford University Press.

Durt, C., Fuchs, T. and Tewes, C. (eds) (2017) *Embodiment, enaction, and culture: investigating the constitution of the shared world*. Cambridge, Massachusetts: MIT Press.

Egginton, W. and Sandbothe, M. (2004) *The Pragmatic Turn in Philosophy: Contemporary Engagements between Analytic and Continental Thought*. State University of New York Press.

Engel, A.K., Friston, K.J. and Kragic, D. (eds) (2015) *The pragmatic turn: toward action-oriented views in cognitive science*. Cambridge, Massachusetts: The MIT Press (Strüngmann forum reports).

Engel, G. (1977) ‘The need for a new medical model: A challenge for biomedicine’, *Science*, 196, pp. 129–136.

Eronen, M.I. (2021) ‘The levels problem in psychopathology’, *Psychological Medicine*, 51(6), pp. 927–933. Available at: https://doi.org/10.1017/S0033291719002514.

Faucher, L. and Forest, D. (eds) (2021) *Defining mental disorder: Jerome Wakefield and his critics*. Cambridge, Massachusetts: The MIT Press (Philosophical psychopathology).

Feest, U. (ed.) (2010) *Historical perspectives on Erklären and Verstehen*. Dordrecht ; New York: Springer (Archimedes, 21).

Feyerabend, P. (1975) *Against Method: Outline of an Anarchistic Theory of Knowledge*. New Left Books.

Frances, A. (2016) ‘Entrenched reductionisms: The bête noire of psychiatry.’, *History of Psychology*, 19(1), pp. 57–59. Available at: https://doi.org/10.1037/hop0000018.

Fried, E.I. and Nesse, R.M. (2015) ‘Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study’, *Journal of Affective Disorders*, 172, pp. 96–102. Available at: https://doi.org/10.1016/j.jad.2014.10.010.

Friesen, P. and Goldstein, J. (2022) ‘Standpoint Theory and the Psy Sciences: Can Marginalization and Critical Engagement Lead to an Epistemic Advantage?’, *Hypatia*, 37(4), pp. 659–687. Available at: https://doi.org/10.1017/hyp.2022.58.

Fuchs, T. (2013) ‘Depression, Intercorporeality, and Interaffectivity Thomas Fuchs Journal of Consciousness Studies 20 (7-8):7-8 (2013’, *Journal of Consciousness Studies*, 20(7–8), pp. 219–238.

Fuchs, T. (2018) *Ecology of the brain: the phenomenology and biology of the embodied mind*. First edition. Oxford, United Kingdom: Oxford University Press.

Fulford, K. (1990) *Moral Theory and Medical Practice*. Cambridge University Press.

Fulford, K.W.M. (2001) ‘“What is (mental) disease?”: an open letter to Christopher Boorse’, *Journal of Medical Ethics*, 27(2), pp. 80–85. Available at: https://doi.org/10.1136/jme.27.2.80.

Fulford, K.W.M. *et al.* (2005) ‘Looking with both eyes open: fact and value in psychiatric diagnosis?’, *World Psychiatry*, 4(2), pp. 78–86.

Fulford, K.W.M. (2008) ‘Values-based practice: A new partner to evidence-based practice and a first for psychiatry’, *Mens Sana Monogr*, 6, pp. 10–21.

Fulford, K.W.M., Bortolotti, L. and Broome, M. (2014) ‘Taking the long view: an emerging framework for translational psychiatric science’, *World Psychiatry*, 13(2), pp. 110–117. Available at: https://doi.org/10.1002/wps.20139.

Fulford, K.W.M. and Handa, A. (2018) 'Categorical and/or continuous? Learning from vascular surgery', *World Psychiatry*, 17(3), pp. 304–305. Available at: https://doi.org/10.1002/wps.20565.

Fulford, K.W.M. and Handa, A. (2021) 'New resources for understanding patients' values in the context of shared clinical decision-making', *World psychiatry: official journal of the World Psychiatric Association (WPA)*, 20(3), pp. 446–447. Available at: https://doi.org/10.1002/wps.20902.

Fulford, K.W.M., Thornton, T. and Graham, G. (2006) *Oxford textbook of philosophy and psychiatry*. Oxford ; New York: Oxford University Press.

Fuss, J. *et al.* (2019) 'Public stakeholders' comments on ICD-11 chapters related to mental and sexual health', *World Psychiatry*, 18(2), pp. 233–235. Available at: https://doi.org/10.1002/wps.20635.

Gagné-Julien, A.-M. (2021a) 'Towards a socially constructed and objective concept of mental disorder', *Synthese*, 198(10), pp. 9401–9426. Available at: https://doi.org/10.1007/s11229-020-02647-7.

Gagné-Julien, A.-M. (2021b) 'Wrongful Medicalization and Epistemic Injustice in Psychiatry: The Case of Premenstrual Dysphoric Disorder', *European journal of analytic philosophy*, 17(2), pp. 5–36. Available at: https://doi.org/10.31820/ejap.17.3.3.

Gallagher, S. (2017) *Enactivist Interventions: Rethinking the Mind*. Oxford University Press.

Gallagher, S. (2018) 'Gallagher S. (2018) Decentering the brain: Embodied cognition and the critique of neurocentrism and narrow-minded philosophy of mind. Constructivist Foundations 14(1): 8–21.', *Constructivist Foundations*, 14(1), pp. 8–21.

Gallagher, S. (2022) 'Integration and Causality in Enactive Approaches to Psychiatry', *Frontiers in Psychiatry*, 13, p. 870122. Available at: https://doi.org/10.3389/fpsyt.2022.870122.

Gallagher, S. and Allen, M. (2018) 'Active inference, enactivism and the hermeneutics of social cognition', *Synthese*, 195(6), pp. 2627–2648. Available at: https://doi.org/10.1007/s11229-016-1269-8.

Gallese, V. and Lakoff, G. (2005) 'The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge', *Cognitive Neuropsychology*, 22, pp. 455–479.

Gallie, W. (1955) 'Essentially contested concepts', *Proceedings of the Aristotelian Society*, 56, pp. 167–198.

Gauld, C. *et al.* (2022) 'From analytic to synthetic-organizational pluralisms: A pluralistic enactive psychiatry', *Frontiers in Psychiatry*, 13, p. 981787. Available at: https://doi.org/10.3389/fpsyt.2022.981787.

Geert, K., Lara, K. and Rico, H. (2017) *Vagueness in Psychiatry*. Oxford University Press.

Ghaemi, S.N. (2007) 'Existence and pluralism: The rediscovery of Karl Jaspers', *Psychopathology*, 40, pp. 75–82.

Ghaemi, S.N. (2009) 'Nosologomania: DSM & Karl Jaspers' Critique of Kraepelin', *Philosophy, Ethics, and Humanities in Medicine*, 4(1), p. 10. Available at: https://doi.org/10.1186/1747-5341-4-10.

Ghaemi, S.N. (2010) *The rise and fall of the biopsychosocial model: reconciling art and science in psychiatry*. Baltimore: Johns Hopkins University Press.

Glackin, S.N. (2016) 'Three Aristotelian Accounts of Disease and Disability: Three Aristotelian Accounts of Disease and Disability', *Journal of Applied Philosophy*, 33(3), pp. 311–326. Available at: https://doi.org/10.1111/japp.12114.

Glackin, S.N. (2017) 'Individualism and the medical: What about somatic externalism?', *Analysis*, 77(2), pp. 287–293. Available at: https://doi.org/10.1093/analys/anx073.

Glackin, S.N. (2019) 'Grounded Disease: Constructing the social from the biological in medicine', *The Philosophical Quarterly*, 69(275), pp. 258–276. Available at: https://doi.org/10.1093/pq/pqy063.

Glackin, S.N. (2020) 'Philosophical Issues in the Addictions', in S. Sussman (ed.) *The Cambridge Handbook of Substance and Behavioral Addictions*. 1st edn. Cambridge University Press, pp. 38–50. Available at: https://doi.org/10.1017/9781108632591.007.

Glackin, S.N., Roberts, T. and Krueger, J. (2021) 'Out of our heads: Addiction and psychiatric externalism', *Behavioral Brain Research*, 398, p. 112936. Available at: https://doi.org/10.1016/j.bbr.2020.112936.

Glennan, S. and Illari, P.M. (2018) *The Routledge handbook of mechanisms and mechanical philosophy*. Abingdon New York (N.Y.): Routledge, Taylor & Francis group (Routledge handbooks in philosophy).

Greenberg, L. (2012) 'Emotions, the great captains of our lives: their role in the process of change in psychotherapy', *American Psychologist*, 67, pp. 697–707.

Guze, S.B. (1992) *Why psychiatry is a branch of medicine*. New York: Oxford University Press.

de Haan, S. *et al.* (2013) 'The phenomenology of deep brain stimulation-induced changes in OCD: an enactive affordance-based model', *Frontiers in Human Neuroscience*, 7. Available at: https://doi.org/10.3389/fnhum.2013.00653.

de Haan, S. (2020) *Enactivism and Psychiatry*. Cambridge University Press.

Hacking, I. (1995) *Rewriting the soul: multiple personality and the sciences of memory*. Princeton, N.J: Princeton University Press.

Haeffel, G.J. *et al.* (2022) 'Folk Classification and Factor Rotations: Whales, Sharks, and the Problems With the Hierarchical Taxonomy of Psychopathology (HiTOP)', *Clinical*

Psychological Science, 10(2), pp. 259–278. Available at: https://doi.org/10.1177/21677026211002500.

Harrington, A. (2019) *Mind Fixers: Psychiatry's Troubled Search for the Biology of Mental Illness*. Norton.

Haslam, N. *et al.* (2020) 'Dimensions over categories: a meta-analysis of taxometric research', *Psychological Medicine*, 50(9), pp. 1418–1432. Available at: https://doi.org/10.1017/S003329172000183X.

Haslam, N. and Kvaale, E.P. (2015) 'Biogenetic Explanations of Mental Disorder: The Mixed-Blessings Model', *Current Directions in Psychological Science*, 24(5), pp. 399–404. Available at: https://doi.org/10.1177/0963721415588082.

Heckers, S. and Kendler, K.S. (2020) 'The evolution of Kraepelin's nosological principles', *World Psychiatry*, 19(3), pp. 381–388. Available at: https://doi.org/10.1002/wps.20774.

Henry, D. (2011) 'Aristotle's pluralistic realism', *The Monist*, 94(2), pp. 197–220.

Hladký, J. and Havlíček, J. (2013) 'Was Tinbergen an Aristotelian? Comparison of Tinbergen's four whys and Aristotle's four causes', *Human Ethology Bulletin*, 28, pp. 3–11.

Horwitz, A.V. and Wakefield, J.C. (2007) *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford University Press.

Huda, A.S. (2019) *The medical model in mental health: an explanation and evaluation*. First edition. Oxford ; New York, NY: Oxford University Press.

Hutto, D.D. and Myin, E. (2013) *Radicalizing enactivism: basic minds without content*. Cambridge, Mass.: MIT Press.

Hyman, S.E. (2010) 'The diagnosis of mental disorders: The problem of reification', *Annual Review of Clinical Psychology*, 6, pp. 155–179.

Hyman, S.E. (2021) 'Psychiatric Disorders: Grounded in Human Biology but Not Natural Kinds', *Perspectives in Biology and Medicine*, 64(1), pp. 6–28. Available at: https://doi.org/10.1353/pbm.2021.0002.

Insel, T.R. and Quirion, R. (2005) 'Psychiatry as a clinical neuroscience discipline', *JAMA*, 294(17), pp. 2221–2224. Available at: https://doi.org/10.1001/jama.294.17.2221.

James, W. (1907) *Pragmatism: A New Name for an Old Way of Thinking*. Longmans, Green and Company.

Jefferson, A. (2022) *Are mental disorders brain disorders?* First edition. Abingdon, Oxon ; New York, NY: Routledge (Routledge focus on philosophy).

Johnson, M. (2017) *Embodied Mind, Meaning, and Reason: How Our Bodies Give Rise to Understanding*. University of Chicago Press.

Kellert, S. *et al.* (2006) *Scientific Pluralism*. University of Minnesota Press.

Kendig, C. (2016) *Natural Kinds and Classification in Scientific Practice*. Routledge.

Kendler, K. and Zachar, P. (2015) 'The incredible insecurity of psychiatric nosology.', in *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*. Baltimore: John Hopkins University Press.

Kendler, K.S. (1990) 'Toward a Scientific Psychiatric Nosology: Strengths and Limitations', *Archives of General Psychiatry*, 47(10), p. 969. Available at: https://doi.org/10.1001/archpsyc.1990.01810220085011.

Kendler, K.S. (2005) 'Toward a philosophical structure for psychiatry', *American Journal of Psychiatry*, 162, pp. 433–440.

Kendler, K.S. (2012) 'The dappled nature of causes of psychiatric illness: replacing the organic–functional/hardware–software dichotomy with empirically based pluralism', *Molecular Psychiatry*, 17(4), pp. 377–388. Available at: https://doi.org/10.1038/mp.2011.182.

Kendler, K.S. (2013) 'What psychiatric genetics has taught us about the nature of psychiatric illness and what is left to learn', *Molecular Psychiatry*, 18, pp. 1058–1066.

Kendler, K.S. (2016a) 'The nature of psychiatric disorders', *World Psychiatry*, 15, pp. 5–12.

Kendler, K.S. (2016b) 'The phenomenology of major depression and the representativeness and nature of DSM criteria', *American Journal of Psychiatry*, 173, pp. 771–780.

Kendler, K.S. (2017) 'DSM disorders and their criteria: how should they inter-relate?', *Psychological Medicine*, 47(12), pp. 2054–2060. Available at: https://doi.org/10.1017/S0033291717000678.

Kendler, K.S. and Campbell, J. (2014) 'Expanding the domain of the understandable in psychiatric illness: an updating of the Jasperian framework of explanation and understanding', *Psychological Medicine*, 44(1), pp. 1–7. Available at: https://doi.org/10.1017/S0033291712003030.

Kendler, K.S., Muñoz, R.A. and Murphy, G. (2010) 'The development of the Feighner Criteria: A historical perspective', *American Journal of Psychiatry*, 167(2), pp. 134–142. Available at: https://doi.org/10.1176/appi.ajp.2009.09081155.

Kendler, K.S., Parnas, J. and Zachar, P. (eds) (2020) *Levels of analysis in psychopathology: cross-disciplinary perspectives*. Cambridge ; New York, NY: Cambridge University Press.

Kendler, K.S., Zachar, P. and Craver, C. (2011) 'What kinds of things are psychiatric disorders?', *Psychological Medicine*, 41, pp. 1143–1150.

Kessler, R.C. (2002) 'The categorical versus dimensional assessment controversy in the sociology of mental illness', *J ournal of Health and Social Behavior*, 43, pp. 171–188.

Kessler, R.C. *et al.* (2003) 'Mild disorders should not be eliminated from the DSM-V', *Archives of General Psychiatry*, 60, pp. 1177–1122.

Kincaid, H. and Sullivan, J. (2014) *Classifying Psychopathology: Mental Kinds and Natural Kinds*. Cambridge: MIT Press.

Kinesella, E. and Pitman, A. (2012) *Phronesis as Professional Knowledge: Practical Wisdom in the Professions*. Sense Publishers.

King, M. and May, J. (eds) (2022) *Agency in mental disorder: philosophical dimensions*. New product. New York: Oxford University Press.

Kirchhoff, M.D. (2015) 'Extended Cognition & the Causal-Constitutive Fallacy: In Search for a Diachronic and Dynamical Conception of Constitution', *Philosophy and Phenomenological Research*, 90(2), pp. 320–360. Available at: https://doi.org/10.1111/phpr.12039.

Kirmayer, L.J. and Ramstead, M. (2017) 'Embodiment and enactment in cultural psychiatry', in *Embodiment, Enaction, and Culture: Investigating the Consitution of the Shared World*.

Klein, D.F. (1964) 'Delineation of two drug-responsive anxiety syndromes', *Psychopharmacologia*, 5(6), pp. 397–408. Available at: https://doi.org/10.1007/BF02193476.

Kleinman (1988) *Rethinking Psychiatry: From Cultural Category to Personal Experience*. Free Press.

Klerman, G. (1978) 'The evolution of a psychiatric nosology', in *Schizophrenia: Science and Practice*. Harvard University Press.

Knox, B. (2022) 'Exclusion of the Psychopathologized and Hermeneutical Ignorance Threaten Objectivity', *Philosophy, Psychiatry, & Psychology*, 29(4), pp. 253–266. Available at: https://doi.org/10.1353/ppp.2022.0044.

Kotov, R. *et al.* (2017) 'The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies.', *Journal of Abnormal Psychology*, 126(4), pp. 454–477. Available at: https://doi.org/10.1037/abn0000258.

Kuhn (1962) *Structure of Scientific Revolutions*. University of Chicago Press.

Kukla, Q.R. (2022) 'What Counts as a Disease, and Why Does It Matter?', *The Journal of Philosophy of Disability*, 2, pp. 130–156. Available at: https://doi.org/10.5840/jpd20226613.

Lakoff, G. and Johnson, M. (2010) *Philosophy in the flesh: the embodied mind and its challenge to Western thought*. Nachdr. New York, NY: Basic Books.

Lakoff and Johnson (1999) *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Basic Books.

LeFrançois, B.A., Menzies, R. and Reaume, G. (eds) (2013) *Mad matters: a critical reader in Canadian mad studies*. Toronto: Canadian Scholars' Press Inc.

Leroi, A.M. (2014) *The Lagoon: how Aristotle invented science*. Translated by S. MacPherson. New York: Viking Pengiun.

Leuridan, B. and Lodewyckx, T. (2021) 'Diachronic causal constitutive relations', *Synthese*, 198(9), pp. 9035–9065. Available at: https://doi.org/10.1007/s11229-020-02616-0.

Lipowsky, Z.J. (1989) 'Psychiatry: Mindless or brainless, both or neither, Canadian Journal of Psychiatry', 34, pp. 249–254.

Longino, H.E. (1990) *Science as social knowledge: values and objectivity in scientific inquiry*. Princeton, N.J: Princeton University Press.

Loughman, A. and Haslam, N. (2018) 'Neuroscientific explanations and the stigma of mental disorder: a meta-analytic study', *Cognitive Research: Principles and Implications*, 3(1), p. 43. Available at: https://doi.org/10.1186/s41235-018-0136-1.

Ludwig, D. and Ruphy, Stephanie (2021) 'Scientific pluralism', in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*.

Luhrmann, T.M. (2001) *Of two minds: an anthropologist looks at American psychiatry*. First Vintage Books edition. New York: Vintage Books.

Maiese, M. (2016) *Embodied Selves and Divided Minds*. Oxford University Press.

Maiese, M. (2021) 'An enactivist reconceptualization of the medical model', *Philosophical Psychology*, 34(7), pp. 962–988. Available at: https://doi.org/10.1080/09515089.2021.1940119.

Maj, M. (2013) 'Mental disorders as "brain diseases" and Jaspers' legacy', *World Psychiatry*, 12(1), pp. 1–3. Available at: https://doi.org/10.1002/wps.20000.

Maj, M. *et al.* (2020) 'The clinical characterization of the adult patient with depression aimed at personalization of management', *World Psychiatry*, 19(3), pp. 269–293. Available at: https://doi.org/10.1002/wps.20771.

Maung, H.H. (2021) 'Causation and Causal Selection in the Biopsychosocial Model of Health and Disease', *European journal of analytic philosophy*, 17(2), pp. 5–27. Available at: https://doi.org/10.31820/ejap.17.2.6.

McHugh, P.R. and Slavney, P.R. (1998) *The perspectives of psychiatry*. 2nd ed. Baltimore: Johns Hopkins University Press.

McLeod, A. (2021) *The Dao of Madness: Mental Illness and Self-Cultivation in Early Chinese Philosophy and Medicine*. 1st edn. Oxford University Press. Available at: https://doi.org/10.1093/oso/9780197505915.001.0001.

Mitchell, S. (2003) *Biological Complexity and Integrative Pluralism*. Cambridge University Press.

Murphy, D. (2006) *Psychiatry in the Scientific Image*. MIT Press.

Nesse, R.M. (2019) *Good Reasons for Bad Feelings: Insights from the Frontier of Evolutionary Psychiatry*. Penguin.

Nesse, R.M. (2019) 'Tinbergen's four questions', *Evolution, Medicine, and Public Health*, 2019(1), pp. 2–2. Available at: https://doi.org/10.1093/emph/eoy035.

Newen, A., De Bruin, L. and Gallagher, S. (2018) *The Oxford Handbook of 4E Cognition*. Oxford University Press.

Nicolson, D.J. (2012) 'The concept of mechanism in biology', *Studies in the History and Philosophy of Biology and Biomedical Sciences*, 43, pp. 152–163.

Nielsen, K. (2021) 'Comparing Two Enactive Perspectives', *Philosophy, Psychiatry, & Psychology*, 28(3), pp. 197–200. Available at: https://doi.org/10.1353/ppp.2021.0031.

Nielsen, K. (2023) *Embodied, embedded, and enactive psychopathology: reimagining mental disorder*. S.l.: PALGRAVE MACMILLAN.

Nielsen, K. and Ward, T. (2020) 'Mental disorder as both natural and normative: Developing the normative dimension of the 3e conceptual framework for psychopathology', *Journal of Theoretical and Philosophical Psychology*, 40(2), pp. 107–123.

Nussbaum (1986) *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge University Press.

Oddli, H.W. *et al.* (2022) 'Causality in psychotherapy research: Towards evidential pluralism', *Psychotherapy Research*, pp. 1–15. Available at: https://doi.org/10.1080/10503307.2022.2161433.

O'Leary, D. (2021) 'Medicine's metaphysical morass: how confusion about dualism threatens public health', *Synthese*, 199(1–2), pp. 1977–2005. Available at: https://doi.org/10.1007/s11229-020-02869-9.

Parfit, D. (2017) *Railton's Defence of Soft Naturalism*. Oxford University Press. Available at: https://doi.org/10.1093/oso/9780198778608.003.0006.

Pearl, J. and Mackenzie, D. (2018) *The Book of Why: The New Science of Cause and Effect*. Basic Books.

Pearson, G. (2018) 'Aristotle on Psychopathology', in P. Kontos (ed.) *Evil in Aristotle*. 1st edn. Cambridge University Press, pp. 122–149. Available at: https://doi.org/10.1017/9781316676813.007.

Pellegrino, E. and Thomasma, D. (1981) *A Philosophical Basis of Medical Practice: Towards a Philosophy and Ethic of the Healing Professions*. Oxford University Press.

Pickard, H. (2017) 'Responsibility without blame for addiction', *Neuroethics*, 10, pp. 169–180.

Pickard, H. and Ahmed, S. (2019) *The Routledge Handbook of Philosophy and Science of Addiction*. Routledge.

Quitkin, F., Rifkin, A. and Klein, D.F. (1979) 'Monoamine Oxidase Inhibitors: A Review of Antidepressant Effectiveness', *Archives of General Psychiatry*, 36(7), p. 749. Available at: https://doi.org/10.1001/archpsyc.1979.01780070027003.

Radden, J. and Sadler, J. (2010) *The Virtuous Psychiatrist: Character Ethics in Psychiatric Practice*. Oxford University Press.

Ramírez-Vizcaya, S. and Froese, T. (2019) 'The Enactive Approach to Habits: New Concepts for the Cognitive Science of Bad Habits and Addiction', *Frontiers in Psychology*, 10, p. 301. Available at: https://doi.org/10.3389/fpsyg.2019.00301.

Ramstead, M.J., Kirchhoff, M.D. and Friston, K.J. (2020) 'A tale of two densities: active inference is enactive inference', *Adaptive Behavior*, 28(4), pp. 225–239.

Reznek, L. (1991) *The Philosophical Defence of Psychiatry*. Routledge.

Richards, R.J. and Daston, L. (eds) (2016) *Kuhn's Structure of Scientific Revolutions at fifty: reflections on a science classic*. Chicago: University of Chicago Press.

de Ridder, J., Peels, R. and van Woudenberg, R. (2018) *Scientism: Prospects and Problems*. Oxford University Press.

Riso, L.P. *et al.* (eds) (2007) *Cognitive schemas and core beliefs in psychological problems: a scientist-practitioner guide*. 1st ed. Washington, DC: American Psychological Association.

Robins, E. and Guze, S.B. (1970) 'Establishment of Diagnostic Validity in Psychiatric Illness: Its Application to Schizophrenia', *American Journal of Psychiatry*, 126(7), pp. 983–987. Available at: https://doi.org/10.1176/ajp.126.7.983.

Rose, D. and Rose, N. (2023) 'Is "another" psychiatry possible?', *Psychological Medicine*, 53(1), pp. 46–54. Available at: https://doi.org/10.1017/S003329172200383X.

Rosqvist, H., Chown, N. and Stenning, A. (eds) (2020) *Neurodiversity studies: a new critical paradigm*. Abingdon, Oxon ; New York, NY: Routledge (Routledge advances in sociology).

Russell, J.L. (2023) 'Problems for enactive psychiatry as a practical framework', *Philosophical Psychology*, pp. 1–24. Available at: https://doi.org/10.1080/09515089.2023.2174423.

Sadler, J. (2004) *Descriptions and Prescriptions: Values, Mental Disorders, and the DSMs*. John Hopkins University Press.

Sadler, J., van Staden, W. and Fulford, K.W.M. (2015) *Oxford Handbook of Psychiatric Ethics*. Oxford University Press.

Sadler, J.Z. (2007) 'The Psychiatric Significance of the Personal Self', *Psychiatry: Interpersonal and Biological Processes*, 70(2), pp. 113–129. Available at: https://doi.org/10.1521/psyc.2007.70.2.113.

Schlick, M. (1985) *General theory of knowledge*. LaSalle, IL: Open Court Pub.

Schramme, T. (2014) 'Christopher Boorse and the philosophy of medicine', *Journal of Medicine and Philosophy*, 39(6), pp. 565–571. Available at: https://doi.org/10.1093/jmp/jhu041.

Schwartz, M. and Wiggins, O. (1985) 'Science, Humanism, and the Nature of Medical Practice: A Phenomenological View', *Perspectives in Biology and Medicine*, 28, pp. 331–361.

Schwartz, M.A., Mokalewicz, M. and Wiggins, O.P. (2017) 'Karl Jaspers: The Icon of Modern Psychiatry', *Israel Journal of Psychiatry and Related Sciences*, 54, pp. 4–7.

Seth, A.K. and Friston, K.J. (2016) 'Active interoceptive inference and the emotional brain', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708), p. 20160007. Available at: https://doi.org/10.1098/rstb.2016.0007.

Shapiro, L. and Spaulding, S. (2021) 'Embodied cognition', in *The Stanford Encylopedia of Philosophy*.

Shapiro, L.A. (2019) *Embodied cognition*. Second Edition. London ; New York: Routledge, Taylor & Francis Group (New problems of philosophy).

Shook, J.R. and Solymosi, T. (2014) *Pragmatist Neurophilosophy: American Philosophy and the Brain*. Bloomsbury Academic.

Sidhu, N.S. *et al.* (2023) 'Competency domains of educators in medical, nursing, and health sciences education: An integrative review', *Medical Teacher*, 45(2), pp. 219–228. Available at: https://doi.org/10.1080/0142159X.2022.2126758.

Slade, M. (2017) 'Implementing shared decision making in routine mental health care', *World psychiatry: official journal of the World Psychiatric Association (WPA)*, 16(2), pp. 146–153. Available at: https://doi.org/10.1002/wps.20412.

Smoller, J.W. *et al.* (2019) 'Psychiatric genetics and the structure of psychopathology', *Molecular Psychiatry*, 24(3), pp. 409–420. Available at: https://doi.org/10.1038/s41380-017-0010-4.

Solomon, M. and Kendler, K.S. (2021) 'The Problem of Aggregating Validators for Psychiatric Disorders', *Journal of Nervous & Mental Disease*, 209(1), pp. 9–12. Available at: https://doi.org/10.1097/NMD.0000000000001256.

Spitzer, R.L. *et al.* (1992) 'Now is the time to retire the term "organic mental disorders"', *American Journal of Psychiatry*, 149, pp. 240–244.

Spitzer, R.L. (1999) 'Harmful dysfunction and the DSM definition of mental disorder.', *Journal of Abnormal Psychology*, 108(3), pp. 430–432. Available at: https://doi.org/10.1037/0021-843X.108.3.430.

Spitzer, R.L. (2001) 'Values and Assumptions in the Development of DSM-III and DSM-III-R: An Insider's Perspective and a Belated Response to Sadler, Hulgus, and Agich's "On Values in

Recent American Psychiatric Classification":', *The Journal of Nervous and Mental Disease*, 189(6), pp. 351–359. Available at: https://doi.org/10.1097/00005053-200106000-00002.

Stein, D.J. (2008) *Philosophy of Psychopharmacology: Smart Pills, Happy Pills, Pepp Pills*. Cambridge University.

Stein, D.J. *et al.* (2010) 'What is a mental/psychiatric disorder? From DSM-IV to DSM-V', *Psychological Medicine*, 40(11), pp. 1759–1765. Available at: https://doi.org/10.1017/S0033291709992261.

Stein, D.J. (2013) 'What is a Mental Disorder? A Perspective from Cognitive-Affective Science', *The Canadian Journal of Psychiatry*, 58(12), pp. 656–662. Available at: https://doi.org/10.1177/070674371305801202.

Stein, D.J. *et al.* (2018) 'Balancing validity, utility, and public health considerations in disorders due to addictive behaviors', *World Psychiatry*, 17, pp. 363–364.

Stein, D.J. (2021) *Problems of living: perspectives from philosophy, psychiatry, and cognitive-affective science*. 1st edn. San Diego: Academic Press is an imprint of Elsevier.

Stein, D.J. *et al.* (2022) 'Psychiatric diagnosis and treatment in the 21st century: paradigm shifts versus incremental integration', *World Psychiatry*, 21(3), pp. 393–414. Available at: https://doi.org/10.1002/wps.20998.

Stein, D.J. (2022) 'Psychiatric disorders are soft natural kinds', *Philosophy, Psychiatry, & Psychology*, 29(3), pp. 183–185.

Stein, D.J., Palk, A.C. and Kendler, K.S. (2021) 'What is a mental disorder? An exemplar-focused approach', *Psychological Medicine*, 51(6), pp. 894–901. Available at: https://doi.org/10.1017/S0033291721001185.

Stein, D.J. and Phillips, K.A. (2013) 'Patient advocacy and DSM-5', *BMC Medicine*, 11(1), p. 133. Available at: https://doi.org/10.1186/1741-7015-11-133.

Stein, D.J. and Reed, G.M. (2019) 'Global mental health and psychiatric nosology: DSM-5, ICD-11, and RDoC', *Brazilian Journal of Psychiatry*, 41(1), pp. 3–4. Available at: https://doi.org/10.1590/1516-4446-2018-4101.

Stoyanov, D. *et al.* (eds) (2021) *International Perspectives in Values-Based Mental Health Practice Case Studies and Commentaries*. Cham: Springer International Publishing.

Strawson, P.F. (1962) 'Freedom and resentment', *Proceedings of the British Academy*, 48, pp. 1–25.

Sullivan, J.A. (2017) 'Coordinated pluralism as a means to facilitate integrative taxonomies of cognition', *Philosophical Explorations*, 20(2), pp. 129–145. Available at: https://doi.org/10.1080/13869795.2017.1312497.

Tabb, K., Lebowitz, M.S. and Appelbaum, P.S. (2019) 'Behavioral Genetics and Attributions of Moral Responsibility', *Behavior Genetics*, 49(2), pp. 128–135. Available at: https://doi.org/10.1007/s10519-018-9916-0.

Tekin, Ş. (2014) 'Self-Insight in the Time of Mood Disorders: After the Diagnosis, Beyond the Treatment', *Philosophy, Psychiatry, & Psychology*, 21(2), pp. 139–155. Available at: https://doi.org/10.1353/ppp.2014.0019.

Tekin, Ş. (2022) 'Participatory Interactive Objectivity in Psychiatry', *Philosophy of Science*, 89(5), pp. 1166–1175. Available at: https://doi.org/10.1017/psa.2022.47.

Tekin, Ş. and Bluhm, R. (eds) (2022) *The Bloomsbury companion to philosophy of psychiatry*. London, UK: Bloomsbury Academic.

Tervalon, M. and Murray-García, J. (1998) 'Cultural Humility Versus Cultural Competence: A Critical Distinction in Defining Physician Training Outcomes in Multicultural Education', *Journal of Health Care for the Poor and Underserved*, 9(2), pp. 117–125. Available at: https://doi.org/10.1353/hpu.2010.0233.

Thompson, E. (2010) *Mind in life: biology, phenomenology, and the sciences of mind*. First Harvard University Press paperback edition. Cambridge, Massachusetts London, England: The Belknap Press of Harvard University Press.

Tinbergen, N. (1963) 'On aims and methods of Ethology', *Zeitschrift für Tierpsychologie*, 20(4), pp. 410–433. Available at: https://doi.org/10.1111/j.1439-0310.1963.tb01161.x.

Tschacher, W., Giersch, A. and Friston, K. (2017) 'Embodiment and Schizophrenia: A Review of Implications and Applications', *Schizophrenia Bulletin*, 43(4), pp. 745–753. Available at: https://doi.org/10.1093/schbul/sbw220.

Tsou, J.Y. (2021) *Philosophy of psychiatry*. Cambridge, United Kingdom: Cambridge University Press.

Varela, F.J., Thompson, E. and Rosch, E. (2016) *The embodied mind: cognitive science and human experience*. revised edition. Cambridge, Massachusetts ; London England: MIT Press.

Wakefield, J.C. (1992) 'Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder', *Psychological Review*, 99, pp. 232–247.

Wakefield, J.C. (2005) 'On winking at the facts, and losing one's Hare: value pluralism and the harmful dysfunction analysis', *World Psychiatry*, 4(2), pp. 88–89.

Wakefield, J.C. (2022) 'Klerman's "credo" reconsidered: NEO-KRAEPELINIANISM , Spitzer's views, and what we can learn from the past', *World Psychiatry*, 21(1), pp. 4–25. Available at: https://doi.org/10.1002/wps.20942.

Wakefield, J.C. and Schmitz, M.F. (2014) 'Uncomplicated depression is normal sadness, not depressive disorder: further evidence from the NESARC', *World Psychiatry*, 13(3), pp. 317–319. Available at: https://doi.org/10.1002/wps.20155.

Ward, D., Silverman, D. and Villalobos, M. (2017) 'Introduction: The Varieties of Enactivism', *Topoi*, 36(3), pp. 365–375. Available at: https://doi.org/10.1007/s11245-017-9484-6.

Whooley, O. (2019) *On the heels of ignorance: psychiatry and the politics of not knowing*. Chicago: The University of Chicago Press.

Williamson, T. (2021) *The Philosophy of Philosophy, Second Edition*. 2nd edn. Hoboken: John Wiley & Sons (The blackwell / brown lectures in philosophy).

Wilson, A.D. and Golonka, S. (2013) 'Embodied Cognition is Not What you Think it is', *Frontiers in Psychology*, 4. Available at: https://doi.org/10.3389/fpsyg.2013.00058.

Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Zachar, P. (2015) 'Psychiatric disorders: natural kinds made by the world or practical kinds made by us?', *World Psychiatry*, 14(3), pp. 288–290. Available at: https://doi.org/10.1002/wps.20240.

Zachar, P. and Kendler, K.S. (2012) 'The removal of pluto from the class of planets and homosexuality from the class of psychiatric disorders: a comparison', *Philosophy, Ethics, and Humanities in Medicine*, 7(1), p. 4. Available at: https://doi.org/10.1186/1747-5341-7-4.

Zachar, P. and Kendler, K.S. (2017) 'The Philosophy of Nosology', *Annual Review of Clinical Psychology*, 13(1), pp. 49–71. Available at: https://doi.org/10.1146/annurev-clinpsy-032816-045020.

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016, Nov 19). An active inference theory of allostasis and interoception in depression. *Philos Trans R Soc Lond B Biol Sci, 371*(1708), 20160011. https://doi.org/10.1098/rstb.2016.0011

Botvinick, M., & Toussaint, M. (2012, Oct). Planning as inference. *Trends in Cognitive Sciences, 16*(10), 485-488. https://doi.org/10.1016/j.tics.2012.08.006

Bruineberg, J., & Rietveld, E. (2014, 08/12

03/31/received

07/17/accepted). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Front Hum Neurosci, 8*, 599. https://doi.org/10.3389/fnhum.2014.00599

Constant, A., Hesp, C., Davey, C. G., Friston, K. J., & Badcock, P. B. (2021). Why Depressed Mood is Adaptive: A Numerical Proof of Principle for an Evolutionary Systems Theory of Depression. *Comput Psychiatr, 5*(1), 60-80. https://doi.org/10.5334/cpsy.70

Duquette, P. (2017, 2017-March-21). Increasing Our Insular World View: Interoception and Psychopathology for Psychotherapists [Review]. *Front Neurosci, 11*(135), 135. https://doi.org/10.3389/fnins.2017.00135

Freund, P., Friston, K., Thompson, A. J., Stephan, K. E., Ashburner, J., Bach, D. R., Nagy, Z., Helms, G., Draganski, B., Mohammadi, S., Schwab, M. E., Curt, A., & Weiskopf, N. (2016, Jun). Embodied neurology: an integrative framework for neurological disorders. *Brain, 139*(Pt 6), 1855-1861. https://doi.org/10.1093/brain/aww076

Friston, K. (2022, 2022/09/02). Computational psychiatry: from synapses to sentience. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-022-01743-z

Friston, K., Mattout, J., & Kilner, J. (2011, Feb). Action understanding and active inference. *Biol Cybern, 104*(1-2), 137-160. https://doi.org/10.1007/s00422-011-0424-z

Kiverstein, J., Rietveld, E., Slagter, H. A., & Denys, D. (2019, May). Obsessive Compulsive Disorder: A Pathology of Self-Confidence? *Trends Cogn Sci, 23*(5), 369-372. https://doi.org/10.1016/j.tics.2019.02.005

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012, Jan). Computational psychiatry. *Trends Cogn Sci, 16*(1), 72-80. https://doi.org/10.1016/j.tics.2011.11.018

Parr, T., & Friston, K. J. (2017, Nov). Uncertainty, epistemics and active inference. *J R Soc Interface, 14*(136), 20170376. https://doi.org/10.1098/rsif.2017.0376

Parr, T., Rees, G., & Friston, K. J. (2018). Computational Neuropsychology and Bayesian Inference. *Front Hum Neurosci, 12*, 61. https://doi.org/10.3389/fnhum.2018.00061

Rae, C. L., Critchley, H. D., & Seth, A. K. (2019, Mar). A Bayesian Account of the Sensory-Motor Interactions Underlying Symptoms of Tourette Syndrome. *Front Psychiatry, 10*, 29, Article 29. https://doi.org/10.3389/fpsyt.2019.00029

Tschacher, W., Giersch, A., & Friston, K. (2017, Jul 1). Embodiment and Schizophrenia: A Review of Implications and Applications. *Schizophr Bull, 43*(4), 745-753. https://doi.org/10.1093/schbul/sbw220